SCHAUM'S ouTlines

全美经典 学习指导系列

统计学原理

（下册）

——推断性统计学

〔美〕S. 伯恩斯坦　R. 伯恩斯坦 著

史道济 译

- 456道有完全解答的习题
- 221道附答案的练习题
- 涵盖本课程的所有基础，是任一教材的补充
- 获取高分的最佳助手

科学出版社



麦格劳-希尔教育出版集团

[O-1541.0101]
责任编辑：刘嘉善

SCHAUM'S ouTlines

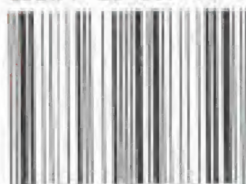
全球销量
超越3000万的
"全美经典学习指导系列"
是您的最佳
学习伴侣!



40年来最畅销的教辅系列
全美著名高校资深教授倾力之作
国内重点高校任课教师全力推荐并担当翻译
省时高效的学习辅导，全面详细的习题解答
迄今为止国内最全面的教辅系列
覆盖大学理工科专业

全美经典 学习指导系列

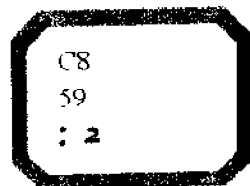
ISBN 7-03-009772-6



9 787030 097729 >

Mc Graw Hill

ISBN 7-03-009772-6/O·1541
定 价：30.00 元



全美经典学习指导系列

统计学原理

（下　册）

——推断性统计学

〔美〕 S. 伯恩斯坦
R. 伯恩斯坦 著

史道济 译

科 学 出 版 社

麦格劳-希尔教育出版集团

2002

内 容 简 介

本书论述工程技术、自然科学和生命科学中常用的统计学原理和方法.全书分为两册.下册主要是推断统计学的方法,包括抽样分布、估计理论、假设检验、回归分析和非参数方法等内容.

本书每一章都有相同的形式.第一部分以大纲的形式论述所有的新概念和新方法.第二部分是各种各样的习题解答,包括许多理论的应用.第三部分是补充的习题,只有答案.这部分内容是检验读者对本书内容的理解程度.

本书内容丰富,叙述严谨.全书有相互参照系统,方便读者阅读.本书适合于高等院校理工科学生、教师和有关的工程技术人员阅读.

Schaum's Outlines

Stephen Bernstein and Ruth Bernstein: Elements of Statistics II: Inferential Statistics

ISBN:0-07-005023-6

图字:01-2001-2118 号

图书在版编目(CIP)数据

统计学原理(下册)——推断性统计学/S. 伯恩斯坦, R. 伯恩斯坦著;史道济译.—北京:科学出版社,2002

(全美经典学习指导系列)

ISBN 7-03-009772-6

Ⅰ.统… Ⅱ.①伯…②伯…③史… Ⅲ.统计学 Ⅳ.C8

中国版本图书馆 CIP 数据核字(2001)第 063366 号

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

http://www.sciencep.com

丽源印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2002 年 1 月第 一 版 开本:A4 (890×1240)

2002 年 1 月第一次印刷 印张:22 1/4

印数:1—5 000 字数:637 000

定价: 30.00 元

(如有印装质量问题,我社负责调换〈环伟〉)

目　录

第十一章　离散型概率分布 ………………………………………………… 1

11.1　离散型概率分布和概率质量函数 ……………………………………… 1

11.2　Bernoulli 试验和多重 Bernoulli 试验 ……………………………………

11.3　二项随机变量，二项试验和二项概率函数……………………………… 2

11.4　二项系数 …………………………………………………………………… 2

11.5　二项概率函数 ……………………………………………………………… 3

11.6　二项概率分布的均值,方差和标准差……………………………………… 4

11.7　二项式展开和二项式定理 ………………………………………………… 5

11.8　Pascal 三角形和二项系数 ………………………………………………… 6

11.9　二项分布族 ………………………………………………………………… 7

11.10　二项累积概率表…………………………………………………………… 7

11.11　批验收抽样………………………………………………………………… 10

11.12　使用方风险和生产方风险………………………………………………… 11

11.13　多元概率分布和联合概率分布…………………………………………… 12

11.14　多项试验…………………………………………………………………… 13

11.15　多项系数…………………………………………………………………… 13

11.16　多项概率函数……………………………………………………………… 14

11.17　多项概率分布族…………………………………………………………… 15

11.18　多项概率分布的均值……………………………………………………… 16

11.19　多项式展开和多项式定理………………………………………………… 16

11.20　超几何试验………………………………………………………………… 16

11.21　超几何概率函数…………………………………………………………… 17

11.22　超几何概率分布族………………………………………………………… 19

11.23　超几何概率分布的均值,方差和标准差 ………………………………… 19

11.24　超几何概率分布的推广…………………………………………………… 20

11.25　超几何分布的二项和多项近似…………………………………………… 20

11.26　Poisson 过程及其随机变量和试验 ……………………………………… 21

11.27　Poisson 概率函数 ………………………………………………………… 22

11.28　Poisson 概率分布族 ……………………………………………………… 22

11.29　Poisson 概率分布的均值,方差和标准差 ………………………………… 23

11.30　Poisson 累积概率表 ……………………………………………………… 24

11.31　Poisson 分布作为二项分布的近似 ……………………………………… 25

第十二章　正态分布和其它连续型概率分布 ………………………………… 36

12.1　连续型概率分布 …………………………………………………………… 36

12.2　正态概率分布和正态概率密度函数 ……………………………………… 37

12.3　正态概率分布族 …………………………………………………………… 38

12.4　正态分布:均值(μ),中位数($\tilde{\mu}$)和众数的关系……………………………… 38

12.5　峰度 ………………………………………………………………………… 39

12.6　标准正态分布 ……………………………………………………………… 39

12.7　标准正态分布和标准正态变量之间的关系 ……………………………… 40

12.8　标准正态分布的面积表 …… 41
12.9　利用 Z 变换计算任意正态分布的概率 …… 41
12.10　单尾概率 …… 43
12.11　双尾概率 …… 45
12.12　二项分布的正态近似 …… 46
12.13　Poisson 分布的正态近似 …… 48
12.14　离散型均匀概率分布 …… 49
12.15　连续型均匀概率分布 …… 50
12.16　指数概率分布 …… 51
12.17　指数分布和 Poisson 分布的关系 …… 51
第十三章　抽样分布 …… 69
13.1　简单随机抽样回顾 …… 69
13.2　独立随机变量 …… 69
13.3　简单随机抽样的数学定义和非数学定义 …… 69
13.4　抽样方法的假定 …… 71
13.5　随机变量 $\overline{X}$ …… 71
13.6　均值的理论抽样分布和经验抽样分布 …… 72
13.7　均值的抽样分布的均值 …… 75
13.8　估计量的准确度 …… 76
13.9　均值的抽样分布的方差:无限总体或有放回抽样 …… 76
13.10　均值的抽样分布的方差:无放回抽样的有限总体 …… 77
13.11　均值的标准误 …… 77
13.12　估计量的精密度 …… 79
13.13　用均值的离散型抽样分布计算概率 …… 79
13.14　用均值的正态抽样分布计算概率 …… 80
13.15　中心极限定理:从有限总体有放回抽样 …… 80
13.16　中心极限定理:从无限总体抽样 …… 83
13.17　中心极限定理:从有限总体无放回抽样 …… 83
13.18　多大是"足够大?" …… 83
13.19　样本和的抽样分布 …… 84
13.20　中心极限定理应用于样本和的抽样分布 …… 85
13.21　二项总体的抽样 …… 85
13.22　成功次数的抽样分布 …… 87
13.23　比率的抽样分布 …… 87
13.24　中心极限定理应用于成功次数的抽样分布 …… 88
13.25　中心极限定理应用于比率的抽样分布 …… 88
13.26　用比率的抽样分布的正态近似计算概率 …… 89
第十四章　总体均值的单样本估计 …… 103
14.1　估计 …… 103
14.2　选择最优估计的标准 …… 103
14.3　均值的估计的标准误 $S_{\bar{x}}$ …… 104
14.4　点估计 …… 104
14.5　点估计的表示和评价 …… 105
14.6　点估计和区间估计的关系 …… 106
14.7　导出 $P(\bar{x}_{1-\alpha/2}\leqslant\overline{X}\leqslant\bar{x}_{\alpha/2})=P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$ …… 106

14.8　导出 $P(\bar{X}-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\mu\leqslant\bar{X}+z_{\alpha/2}\sigma_{\bar{x}})=1-\alpha$ …… 107
14.9　总体均值 μ 的置信区间:标准差 σ 已知的正态分布总体 …… 108
14.10　置信限的表示 …… 109
14.11　置信区间的精度 …… 109
14.12　已知标准差确定样本容量 …… 111
14.13　总体均值 μ 的置信区间:来自标准差 σ 已知的任何总体的大样本($n\geqslant30$) …… 112
14.14　确定总体均值 μ 的置信区间:总体标准差 σ 未知 …… 112
14.15　t 分布 …… 113
14.16　t 分布和标准正态分布的关系 …… 114
14.17　自由度 …… 114
14.18　术语"Student t 分布" …… 115
14.19　t 分布的临界值 …… 115
14.20　表 A.6:t 分布的临界值 …… 117
14.21　总体均值 μ 的置信区间:来自标准差 σ 未知的正态总体的小样本($n<30$) …… 118
14.22　确定样本容量:来自标准差 σ 未知的正态总体的小样本 …… 120
14.23　总体均值 μ 的置信区间:来自标准差 σ 未知的正态总体的大样本($n\geqslant30$) …… 121
14.24　总体均值 μ 的置信区间:来自标准差 σ 未知的非正态总体的大样本($n\geqslant30$) …… 122
14.25　总体均值 μ 的置信区间:来自非正态总体的小样本($n<30$) …… 122
第十五章　总体方差、标准差及比率的单样本估计 …… 133
15.1　方差、标准差及比率的最优估计 …… 133
15.2　χ^2 统计量和 χ^2 分布 …… 133
15.3　χ^2 分布的临界值 …… 134
15.4　表 A.7:χ^2 分布的临界值 …… 135
15.5　正态分布总体方差 σ^2 的置信区间 …… 136
15.6　置信限的表示 …… 137
15.7　方差置信区间的精度 …… 138
15.8　确定为得到方差的所要求估计性质所必需的样本容量 …… 139
15.9　用近似正态方法确定方差的置信区间 …… 139
15.10　用样本方差的抽样分布来近似总体方差的置信区间 …… 140
15.11　正态分布总体标准差 σ 的置信区间 …… 141
15.12　用样本标准差的抽样分布来近似总体标准差的置信区间 …… 142
15.13　二项总体比率 p 的最优估计 …… 142
15.14　二项总体比率 p 的近似置信区间的导出 …… 142
15.15　参数 p 的估计 …… 144
15.16　当 p 未知时,确定何时 n 为"足够大" …… 144
15.17　有限总体无放回抽样对二项参数 p 的近似置信区间 …… 145
15.18　二项参数 p 的精确的置信区间 …… 145
15.19　二项参数 p 的近似置信区间估计的精度 …… 145
15.20　确定二项参数 p 近似置信区间的样本容量 …… 146
15.21　二项总体百分比的近似置信区间 …… 147
15.22　二项总体总的成功次数的近似置信区间 …… 147
15.23　估计总体容量 N 的捕获-再捕获方法 …… 147

第十六章　单样本的假设检验 …… 157
16.1　统计假设检验 …… 157
16.2　零假设和对立假设 …… 157
16.3　零假设的检验 …… 158
16.4　单侧与双侧假设检验 …… 158
16.5　总体均值 μ 的假设检验:标准差 σ 已知的正态分布总体 …… 158
16.6　P 值 …… 159
16.7　第Ⅰ类错误和第Ⅱ类错误 …… 160
16.8　临界值与临界域 …… 160
16.9　显著水平 …… 162
16.10　统计假设检验的决策规则 …… 163
16.11　统计假设的选择 …… 164
16.12　第Ⅱ类错误概率 …… 164
16.13　使用方风险和生产方风险 …… 165
16.14　为何不能证明零假设 …… 165
16.15　古典推断与 Bayes 推断 …… 165
16.16　检验零假设的步骤 …… 166
16.17　用 $\overline{X}$ 作为检验统计量的假设检验 …… 167
16.18　检验的功效,操作特性曲线和功效曲线 …… 168
16.19　总体均值 μ 的假设检验:取自未知标准差 σ 的正态分布总体的小样本($n<30$) …… 169
16.20　t 统计量的 P 值 …… 169
16.21　t 统计量的假设检验决策规则 …… 170
16.22　β,$1-\beta$,功效曲线和 OC 曲线 …… 171
16.23　总体均值 μ 的假设检验:来自任意分布总体的大样本($n\geqslant30$) …… 171
16.24　单样本参数假设检验的假定条件 …… 171
16.25　违背假定的情况 …… 172
16.26　检验正态分布总体方差 σ^2 的假设 …… 172
16.27　检验正态分布总体标准差 σ 的假设 …… 174
16.28　检验二项总体比率 p 的假设:大样本 …… 174
16.29　检验二项总体比率 p 的假设:小样本 …… 175
第十七章　两样本估计和假设检验 …… 189
17.1　独立样本和成对样本 …… 189
17.2　两总体均值差($\mu_1-\mu_2$)的最优估计 …… 189
17.3　均值差的理论抽样分布 …… 190
17.4　均值差($\mu_1-\mu_2$)的置信区间:标准差(σ_1 和 σ_2)已知的正态分布总体的独立样本 …… 190
17.5　均值差($\mu_1-\mu_2$)的假设检验:标准差(σ_1、σ_2)已知的正态分布总体的独立样本 …… 191
17.6　两均值差的标准误的估计 …… 192
17.7　均值差($\mu_1-\mu_2$)的置信区间:标准差未知,但假定相等($\sigma_1=\sigma_2$)的正态分布总体的独立小样本($n_1<30$ 且 $n_2<30$) …… 193
17.8　均值差($\mu_1-\mu_2$)的假设检验:标准差未知,但假设相等($\sigma_1=\sigma_2$)的正态分布总体的独立小样本($n_1<30$ 且 $n_2<30$) …… 194

17.9　均值差($\mu_1-\mu_2$)的置信区间:标准差(σ_1 和 σ_2)未知的任何总体分布的独立大样本($n_1\geqslant30$ 且 $n_2\geqslant30$) …… 195

17.10　均值差($\mu_1-\mu_2$)的假设检验:标准差(σ_1 和 σ_2)未知的任意分布总体的独立大样本($n_1\geqslant30$ 且 $n_2\geqslant30$) …… 195

17.11　均值差($\mu_1-\mu_2$)的置信区间:成对样本 …… 195

17.12　均值差($\mu_1-\mu_2$)的假设检验:成对样本 …… 199

17.13　均值的两样本参数估计和假设检验的假定 …… 200

17.14　如果违背假定 …… 201

17.15　独立样本和成对样本方法在精确性和功效方面的比较 …… 201

17.16　F 统计量 …… 201

17.17　F 分布 …… 202

17.18　F 分布的临界值 …… 203

17.19　表 A.8:F 分布的临界值 …… 205

17.20　方差比(σ_1^2/σ_2^2)的置信区间:参数($\sigma_1^2,\sigma_1,\mu_1$ 和 $\sigma_2^2,\sigma_2,\mu_2$)未知的正态分布总体的独立样本 …… 206

17.21　方差比(σ_1^2/σ_2^2)的假设检验:参数($\sigma_1^2,\sigma_1,\mu_1$ 和 $\sigma_2^2,\sigma_2,\mu_2$)未知的正态分布总体的独立样本 …… 208

17.22　何时检验方差齐性 …… 209

17.23　比率差(p_1-p_2)的最优估计量:独立大样本 …… 210

17.24　比率差的理论抽样分布 …… 210

17.25　二项总体比率差(p_1-p_2)的近似置信区间:独立大样本 …… 211

17.26　两个二项总体比率差(p_1-p_2)的假设检验:独立大样本 …… 212

第十八章　多个样本的参数估计与假设检验 …… 227

18.1　多个样本推断 …… 227

18.2　方差分析 …… 227

18.3　单向、双向及多向方差分析 …… 227

18.4　单向方差分析:固定效应、随机效应 …… 228

18.5　单向固定效应方差分析:各种假定 …… 228

18.6　样本容量相等时的单向固定效应方差分析:H_0 与 H_1 …… 228

18.7　样本容量相等时的单向固定效应方差分析:数据的整理 …… 229

18.8　样本容量相等时的单向固定效应方差分析:基本原理 …… 230

18.9　$SST=SSA+SSW$ …… 230

18.10　SST 与 SSA 的计算公式 …… 231

18.11　自由度与均方 …… 232

18.12　F 检验 …… 233

18.13　方差分析表 …… 234

18.14　多重比较检验 …… 235

18.15　Duncan 多重极差检验 …… 235

18.16　多重比较的后继问题:置信区间的计算 …… 236

18.17　方差齐性的检验 …… 238

18.18　单向固定效应方差分析:样本容量相等或不等 …… 239

18.19　一般程序的单向固定效应方差分析:数据的整理 …… 239

18.20　一般程序的单向固定效应方差分析:平方和 …… 240

18.21　一般程序的单向固定效应方差分析:自由度与均方 …… 240

18.22　一般程序的单向固定效应方差分析:F 检验 …… 241
18.23　一般程序的单向固定效应方差分析:多重比较 …… 242
18.24　一般程序的单向固定效应方差分析:置信区间的计算及方差齐性的检验 …… 243
18.25　方差分析前提假定的破坏 …… 243
第十九章　回归和相关 …… 255
19.1　两变量间关系的研究 …… 255
19.2　简单线性回归模型 …… 255
19.3　最小二乘回归直线 …… 256
19.4　方差 $\sigma^2_{Y\cdot X}$ 的估计 …… 258
19.5　截距 $\hat{a}$ 和斜率 $\hat{b}$ 的均值和方差 …… 259
19.6　截距 a 和斜率 b 的置信区间 …… 260
19.7　方差 $\sigma^2_{Y\cdot X}$ 的置信区间 …… 261
19.8　Y 期望值的预测区间 …… 262
19.9　关于斜率 b 的假设检验 …… 262
19.10　两样本或多样本的简单线性回归方程的比较 …… 263
19.11　多元线性回归 …… 263
19.12　简单线性相关 …… 264
19.13　相关系数 r 的导出 …… 265
19.14　总体相关系数 ρ 的置信区间 …… 267
19.15　用 r 分布检验总体相关系数 ρ 的假设 …… 268
19.16　用 t 分布检验关于 ρ 的假设 …… 269
19.17　用 Z 分布检验假设 $\rho = c$ …… 269
19.18　样本相关系数 r 的解释 …… 270
19.19　多重相关与偏相关 …… 270
第二十章　非参数方法 …… 289
20.1　非参数方法与参数方法 …… 289
20.2　χ^2 检验 …… 289
20.3　χ^2 拟合优度检验 …… 289
20.4　独立性的 χ^2 检验:列联表分析 …… 290
20.5　k 个二项比率齐性的 χ^2 检验 …… 292
20.6　秩次检验 …… 293
20.7　单样本检验:Wilcoxon 符号秩检验 …… 293
20.8　两样本检验:相依样本的 Wilcoxon 符号秩检验 …… 295
20.9　两样本检验:独立样本的 Mann-Whitney U 检验 …… 296
20.10　多样本检验:k 个独立样本的 Kruskal-Wallis H 检验 …… 299
20.11　Spearman 秩相关检验 …… 301
附录 …… 325
表 A.3　累积二项概率 …… 325
表 A.4　累积 Poisson 概率 …… 327
表 A.5　标准正态分布的面积 …… 328
表 A.6　t 分布的临界值 …… 329
表 A.7　χ^2 分布的临界值 …… 330
表 A.8　F 分布的临界值 …… 331
表 A.9　最小显著的学生化极差 r_p …… 339

表 A. 10　r 到 Z_r 的变换 …… 340
表 A. 11　Pearson 乘积矩相关系数 r 的临界值 …… 342
表 A. 12　Wilcoxon W 的临界值 …… 343
表 A. 13　Mann-Whitney U 的临界值 …… 344
表 A. 14　Kruskal-Wallis H 的临界值 …… 345
表 A. 15　Spearman r_s 的临界值 …… 346

第十一章　离散型概率分布

11.1　离散型概率分布和概率质量函数

在上册第十章我们已经讲过离散型概率分布的一般性质.当时(见上册,10.3节)指出,离散型概率分布就是在离散随机变量定义的样本空间中指定事件发生概率的概率函数.当离散随机变量在定义域中取任意值($X=x$)时,概率函数对这个值就有指定的概率[$P(X=x)=f(x)$].离散型概率分布的概率函数称为概率质量函数,因为概率是散布在随机变量的各离散值上.

当所有的离散型概率分布都用概率质量函数的唯一确定的公式定义时,它们可以表示为以下4种形式:函数本身,由函数计算得到的概率分布列,概率表或概率图(见上册,表10.1和图10-1).虽然,对数学家而言,术语"离散型概率分布"和"概率质量函数"是同义词,但是,我们将区别离散型概率分布的定义函数和用该函数计算出的概率值的理论分布.

本章将讨论4个最重要的离散型概率分布:**二项分布**,**多项分布**,**超几何分布**和 **Poisson 分布**的具体特征和应用.先从二项分布开始.

11.2　Bernoulli 试验和多重 Bernoulli 试验

要理解二项分布,必须首先理解两个基本概念:**Bernoulli 试验**和**多重 Bernoulli 试验**;它们以瑞士数学家 James Bernoulli(也称 Jakob Bernoulli)(1645－1705)的名字命名,Bernoulli 首先研究了二者的性质.Bernoulli 对概率论的其它贡献是提出了 Bernoulli 定理(见上册,8.2节).Bernoulli 试验和多重 Bernoulli 试验有以下性质:

(1) 一次 Bernoulli 试验有且仅有两种可能结果,称为"成功"和"失败".两个结果是随机决定且互斥的.

(2) 多重 Bernoulli 试验是同一Bernoulli 试验重复 n 次的序列.

(3) 在每次试验中,成功的概率是 p,失败的概率是 $q=1-p$.

(4) 各次 Bernoulli 试验之间**相互独立**(见上册,9.4节):每次试验结果不受其它各次试验结果的影响.

(5) 对每次 Bernoulli 试验,成功的概率 p 均相同(为**常数**)(这表明失败的概率 q 也是常数).

例 11.1　一只碗中放有20粒除颜色外别无差异的石子:10粒为红色,10粒为绿色.以下哪个是多重 Bernoulli 试验:(a)闭上眼睛相继取出10粒石子,每次记录石子是否为红色,然后将石子放回,(b)重复(a)的试验,但每次取出的石子不放回?

解 ☞　(a)是多重 Bernoulli 试验.这里,对每一次试验:成功=红色,失败=绿色,$p=10/20=\frac{1}{2}$,$q=1-\frac{1}{2}=\frac{1}{2}$,并且各次试验相互独立.

(b)不是多重 Bernoulli 试验.因为取出的石子没有重新放回,p 在各次试验中不是保持不变的,且各次试验相互不独立(一次试验的结果对随后的试验结果有影响).

在例11.1(b)描述的试验中,假设碗中放有150粒红色石子和150粒绿色石子,而不是红色和绿色各10粒,再次进行试验,即闭上眼睛相继取出10粒石子,每次记录石子是否为红色,并且石子取出后不放回.试验似乎同例11.1(b)是一样的,但是,对统计学家而言,它们差异很

大,因为总体容量改变了[1]. 对于 300 粒石子,样本容量 n 对总体容量 N 的比例是 $10/300=0.033$,而对于 20 粒石子,比例是 $10/20=0.50$. 一般,统计学家认为,在抽样条件下,如果样本容量对总体容量的比例不大于 0.05(即,样本不大于总体的 5%,或 $n\leqslant 0.05N$),对不太保守的统计学家而言不超过 0.10($n\leqslant 0.10N$),就可以假设 p"基本上"保持不变,各次试验"基本上"相互独立;如果多重 Bernoulli 试验要求的所有其它性质均能满足,就可以认为抽样是多重 Bernoulli 试验.

我们在这里处理的问题以及之前和今后还将多次处理的问题是:纯粹的抽象统计模型和待解决的实际问题之间的冲突. 统计方法建立在理论的,理想化环境的数学模型的基础之上,但这些模型在现实世界中极少存在. 因此,一个统计模型的严格假设和要求往往不能完全满足. 但是,在许多实际情形,就像这里,如果假设和要求可以"基本上"满足,则给定的统计方法能够给出合理的准确结果.

多重 Bernoulli 试验的假设仅在以下两种情形可完全满足:(1) **从有限或无限总体中有放回随机抽样**(见上册,3.16 节),(2) **从无限总体中无放回抽样**. 在第一种情形,总体可以较小也可以相对于样本是较大的,由于每次将抽样个体放回,p 为常数且各次试验相互独立. 在第二种情形,总体必须足够大,使得不放回抽样个体不会对实际问题造成影响;这样,p 仍为常数且各次试验相互独立.

11.3 二项随机变量,二项试验和二项概率函数

如果随机变量 X 用作记录 n 重 Bernoulli 试验中出现的成功次数($X=x$),则 X 称为**二项随机变量**,n 重 Bernoulli 试验序列称为一个**二项**(binomial)**试验**(bi,两个;nomial,项或结果). 这样,如果一个概率函数被用作给出由二项随机变量定义的样本空间的每一样本点的概率值,则该函数称为**二项概率质量函数**或**二项概率函数**或**二项概率分布**或**二项分布**.

例 11.2 由上面和 11.2 节的定义说明:掷一枚硬币 3 次,出现正面的次数的离散型概率分布是一个二项分布.

解 ☞ 二项分布必须有以下基本要素:Bernoulli 试验,多重 Bernoulli 试验,二项随机变量,二项试验以及对二项变量定义的样本空间中每一样本点指定概率值的概率函数.

在这个问题中,Bernoulli 试验是掷一枚硬币,它只有两个可能结果:正面和反面,而且两个结果是随机决定且互斥的. 可以将其中一个结果认定为成功结果,这种选择依赖于所研究的实际概率问题. 这里,将观测到正面视为成功,目的是确定出现各种正面次数的概率. 应该注意的是:将某一结果视为成功并不意味着该结果更能令人接受;它只表明,该结果是所要研究的. 在任何情形,将正面视为成功从而反面视为失败之后,Bernoulli 试验就可以定义为:成功(正面)的概率是 $p=\frac{1}{2}$, 同时,失败(反面)的概率是 $q=1-\frac{1}{2}=\frac{1}{2}$.

n 重 Bernoulli 试验序列是掷一枚硬币 3 次,每掷一次就是一次 Bernoulli 试验. 在这个序列中,对每次试验均成立:(1) 只有同样的两个互斥且随机决定的可能结果:成功(正面)或失败(反面),(2) 对各次试验,成功的概率恒为常数 $p=\frac{1}{2}$, (3) 各次试验相互独立[假定它是理想化的游戏(见上册,8.1 节),硬币的任何结果都不影响随后的投掷结果].

离散随机变量 X 记录 $n=3$ 重 Bernoulli 试验中出现成功的次数($X=x$),即出现正面的不同次数. X 是一个二项随机变量,这使得多重 Bernoulli 试验序列成为一个二项试验. 最后,一个概率函数被用来(见上册,例 10.5)指定由二项随机变量定义的样本空间 $S=\{0,1,2,3\}$ 的每一样本点的概率,从而,得到一个二项概率分布(见上册,表 10.2).

11.4 二项系数

上册第九章介绍了三个计数规则:乘法原理(见上册,9.12 节),排列(见上册,9.13 节)和组合(见上册,9.14 节). **二项系数**也是一个计数规则,它是二项概率函数(见下面 11.5 节)的

[1] 原文为样本容量改变了.——译者注

一个组成部分.如果 n 个个体只由不同的两类构成,一类有 x 个,另一类有 $n-x$ 个,则二项系数用于确定这 n 个个体构成的可能排列数.可以证明,这个排列数由 n 个不同个体中一次选取 x 个的组合数公式[上册,公式(9.20)]给出

$$\text{二项系数} = {}_nC_x = \binom{n}{x} = \frac{n!}{x!(n-x)!} \tag{11.1}$$

为理解上述含义,我们用 4 粒石子,红色(记为 R)和绿色(记为 G)各有 2 粒,试问:将这 $n=4$个个体**按次序**排成一列共有多少种排法? 由于我们考虑个体的次序,故这显然是排列问题.如果 4 个个体全不相同,则答案是(见上册,习题 9.30)

$${}_nP_n = n! = 4! = 24$$

如果 4 个个体全不相同,但不考虑次序,即仅仅关心 n 个个体中同时取出 4 个的可能组合数,则答案是(见上册,习题 9.41)

$${}_nC_n = 1$$

当 $n=4$ 且有($x=2$)个为红色和($n-x=2$)个为绿色时,两个答案都不正确:将 4 粒石子实际排成一列,可以看出,可能的排法是

$$RRGG \quad RGRG \quad RGGR \quad GGRR \quad GRGR \quad GRRG$$

这里一类有 $x=2$ 个,另一类有 $n-x=2$ 个,($n=4$)个个体只有 6 种可能的排列.我们可用二项系数给出这个答案:

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{4!}{2!2!} = 6$$

例 11.3 一只碗中放有 20 粒除颜色外别无差异的石子,12 粒为红色,8 粒为绿色.闭上眼睛从碗中取出 6 粒石子,观测每粒石子的颜色,然后将其放回.试验中,有多少种取出 4 粒红色石子的方法?

解 在这个二项试验中,$n=6$ 且 $x=4$.由二项系数得

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} = \frac{6!}{4!2!} = 15$$

11.5 二项概率函数

用上册例 10.5 中的概率函数计算二项分布 掷一枚硬币 3 次出现正面的次数 的概率值,是基于对 k 个独立事件的特殊乘法原理的推广[上册,公式(9.9)]和 k 个互斥事件的特殊加法原理的推广(上册,8.6 节,性质 4).现在,我们给出一个公式,称为**二项概率函数**,它将这些乘法和加法合成一个计算公式

$$f(x) = \binom{n}{x} p^x q^{n-x}, \; x = 0, 1, 2, \cdots, n \tag{11.2}$$

这说明,在一个二项试验的 n 重 Bernoulli 试验中,x 次成功的概率[$f(x)$]等于二项系数[公式(11.1)]和成功概率 p 自乘 x 次(p^x)(见上册,1.9 节)以及失败概率 q 自乘 $n-x$ 次 q^{n-x} 的乘积.

例 11.4 如例 11.3,一只碗中放有 12 粒红色石子和 8 粒绿色石子,仍取出 6 粒,不同的是一粒一粒地

取,每次观测颜色并将石子放回.试问:在 6 次选取中,取出 4 粒红色的概率是多少?

解 从例 11.3 知,二项系数 $\binom{n}{r}$ 是 15;即取得一个容量为 6 的样本使之含有 4 粒红色石子的方法有 15 种.因此,为求得取出 4 粒红色的概率[$P(X=4)=f(4)$],只需将取出 4 粒红色石子和 2 粒绿色石子的概率乘以这个系数.对每次试验,取得一粒红色的概率(p)是 $\frac{12}{20}=0.6$,取得一粒绿色的概率(q)是 $1-p=0.4$.4 次试验每次均取得一粒红色的概率是$(0.6)(0.6)(0.6)(0.6)$,或 0.6^4;2 次试验每次均取得一粒绿色的概率是$(0.4)(0.4)$,或 0.4^2.由公式(11.2),取得一个容量为 6 且含有 4 粒红色石子的样本的概率是

$$f(x)=\binom{n}{x}p^x q^{n-x}$$

$$f(4)=(15)(0.6)^4(0.4)^2=0.311$$

11.6 二项概率分布的均值,方差和标准差

任意离散型概率分布都有均值,方差和标准差.由上册公式(10.10),离散随机变量的概率分布的期望(或均值)的一般公式为

$$E(X)=\mu=\sum_x xf(x)$$

为得到**二项概率分布的均值**的公式,只须简单的以公式(11.2)替换 $f(x)$:

$$E(X)=\mu=\sum_x x\binom{n}{x}p^x q^{n-x} \tag{11.3}$$

在数学上可以证明,上式可简化为

$$E(X)=\mu=np \tag{11.4}$$

由上册公式(10.22),离散型概率分布的方差的一般计算公式为

$$\sigma^2=\sum_x x^2 f(x)-\mu^2$$

用公式(11.2)替换 $f(x)$,公式(11.4)替换 μ,得到**二项概率分布的方差**的公式:

$$\sigma^2=\sum_x x^2\left[\binom{n}{x}p^x q^{n-x}\right]-(np)^2 \tag{11.5}$$

在数学上可以证明,上式可简化为

$$\sigma^2=npq \tag{11.6}$$

离散型概率分布的标准差是

$$\sigma=\sqrt{\sigma^2}$$

从而**二项概率分布的标准差**是

$$\sigma=\sqrt{npq} \tag{11.7}$$

例 11.5 求二项概率分布(掷一枚硬币 3 次,出现正面的次数)的均值,方差和标准差.

解 对于这个试验,p=掷一次硬币出现正面的概率$=\frac{1}{2}$,q=掷一次硬币出现反面的概率$=\frac{1}{2}$,$n=3$.将 n 和 p 的值代入公式(11.4),

$$E(X) = \mu = np = 3\left(\frac{1}{2}\right) = 1.5$$

将 n、p 和 q 的值代入(11.6)，

$$\sigma^2 = npq = (3)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 0.75$$

取方差的平方根可以计算标准差.

$$\sigma = \sqrt{npq} = \sqrt{0.75} = 0.87$$

11.7 二项式展开和二项式定理

二项概率函数和二项分布的名称来源于它们与**二项式展开**的关系.形如$(a+b)$的**二项代数表达式**(见上册,1.11 节)自乘 n 次后产生的**若干项和**,称为二项式展开.下面是二项式展开的 2 个例子:

$$(a+b)^2 = a^2 + 2ab + b^2 \quad (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

二项式定理是如下展开$(a+b)^n$ 的一般公式,这里 a 和b 是实数,n 和x 是正整数

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^{n-x} b^x \tag{11.8}$$

它是(1)二项系数[公式(11.1)],(2)a 自乘$(n-x)$次(a^{n-x}),和(3)b 自乘x 次(b^x)的乘积从 $x=0$ 到 $x=n$ 的和.例如,二项表达式$(a+b)^2$ 的二项式展开是

$$\begin{aligned}(a+b)^2 &= \sum_{x=0}^{2} \binom{2}{x} a^{2-x} b^x \\ &= \binom{2}{0} a^{2-0} b^0 + \binom{2}{1} a^{2-1} b^1 + \binom{2}{2} a^{2-2} b^2 \\ &= \frac{2!}{0!2!} a^2 + \frac{2!}{1!1!} ab + \frac{2!}{2!0!} b^2 \\ &= a^2 + 2ab + b^2\end{aligned}$$

如果设 $a=q$ 且$b=p$,这里 q 和p 分别是一次 Bernoulli 试验的失败和成功的概率,n = 多重 Bernoulli 试验的次数,x = 成功次数,则

$$\begin{aligned}(q+p)^n &= \sum_{x=0}^{n} \binom{n}{x} q^{n-x} p^x \\ &= \binom{n}{0} q^{n-0} p^0 + \binom{n}{1} q^{n-1} p^1 + \cdots + \binom{n}{n-1} q^{n-(n-1)} p^{n-1} + \binom{n}{n} q^{n-n} p^n \\ &= q^n + nq^{n-1}p + \cdots + nqp^{n-1} + p^n \end{aligned} \tag{11.9}$$

(注:这里颠倒 p 和q 的次序,是为了与二项式定理中两个参数的次序保持一致,其中 a^{n-x}先于b^x). 从上面展开式可以看到,二项概率函数和二项式展开$(q+p)^n$ 之间有关系:对每一个整数 x,x 次成功的概率对应二项式展开的一项.

例 11.6 用二项式定理确定二项变量(掷一粒骰子 5 次得到 5 点的次数)的概率分布.

表 11.1

5 点的次数 x	概率 $f(x)=\binom{n}{x}p^xq^{n-x}$
0	$f(0)=\binom{5}{0}(1/6)^0(5/6)^{5-0}=0.40187$
1	$f(1)=\binom{5}{1}(1/6)^1(5/6)^{5-1}=0.40188$
2	$f(2)=\binom{5}{2}(1/6)^2(5/6)^{5-2}=0.16076$
3	$f(3)=\binom{5}{3}(1/6)^3(5/6)^{5-3}=0.03215$
4	$f(4)=\binom{5}{4}(1/6)^4(5/6)^{5-4}=0.00322$
5	$f(5)=\binom{5}{5}(1/6)^5(5/6)^{5-5}=0.00013$
$\sum$	1.00001

解 ☞ 由公式(11.9),当 $n=5$ 时

$$
\begin{aligned}
(q+p)^5 &= \sum_{x=0}^{5}\binom{5}{x}q^{5-x}p^x \\
&= \binom{5}{0}q^{5-0}p^0+\binom{5}{1}q^{5-1}p^1+\binom{5}{2}q^{5-2}p^2 \\
&\quad +\binom{5}{3}q^{5-3}p^3+\binom{5}{4}q^{5-4}p^4+\binom{5}{5}q^{5-5}p^5 \\
&= \frac{5!}{0!5!}q^5+\frac{5!}{1!4!}q^4p+\frac{5!}{2!3!}q^3p^2+\frac{5!}{3!2!}q^2p^3+\frac{5!}{4!1!}qp^4+\frac{5!}{5!0!}p^5 \\
&= q^5+5q^4p+10q^3p^2+10q^2p^3+5qp^4+p^5
\end{aligned}
$$

如果将这个展开式与表 11.1 中变量的概率分布相比较,你会发现,任一 x 次成功的概率对应着展开式的一项.

由于一次 Bernoulli 试验只有成功和失败两个可能结果,并且它们之间互斥,

$$q+p=1$$

因此

$$(q+p)^n=\sum_{x=0}^{n}\binom{n}{x}q^{n-x}p^x=1$$

表 11.1 表明二项概率函数和二项式定理之间的这种关系:即对于离散型概率分布,有(见上册,图 10-3,性质 7)

$$\sum_x f(x)=\sum_x\binom{n}{x}p^xq^{n-x}=1$$

11.8 Pascal 三角形和二项系数

Pascal 三角形[①]是$(a+b)^n$ 的展开式中二项系数的一个排列. $n=5$ 的 Pascal 三角形如图 11-1 所示. 可以看出,每一行的第一个数字和最后一个数字都是 1,并且一行中间的任意数字

① 我国南宋时期数学家杨辉在他所著的《详解九章算术》中记载着有关二项系数的研究,Pascal 三角形与此类似,在我国也称杨辉三角形.——译者注

都能通过对该数字上方处于它左右两边的两个数字求和计算得到. 例如,第 $n=5$ 行的两个 5 就是它们上方处于左右两边的数字 1 和 4 的和. 在 $n=5$ 行,$(a+b)^5$[或$(q+p)^5$]的系数是:1,5,10,10,5,1. 显然,这些与我们在例 11.6 得到的$(q+p)^5$ 的系数相同.

n	
0	1
1	1 1
2	1 2 1
3	1 3 3 1
4	1 4 6 4 1
5	1 5 10 10 5 1

图 11-1

例 11.7 用 Pascal 三角形求$(q+p)^6$ 的展开式的二项系数.

解 ☞ 用上面的计算方法,系数是: 1,6,15,20,15,6,1.

11.9 二项分布族

二项概率函数只有一种形式[公式(11.2)]

$$f(x)=\binom{n}{x}p^x q^{n-x}$$

但是它定义了无数多个具体的二项分布,对常数 n 和 p 的不同组合即为一个. 指定 q 是没有必要的,因为 $q=1-p$. 由于指定 n 和 p 即完全决定了所考虑的是哪个具体的二项分布,故 n 和 p 称为**二项分布的参数**. 注意:在这里,术语"参数"的使用不同于在上册 3.4 节中介绍的描述性总体参数-由总体计算的描述性测量值. 尽管理论方程的参数是确定常数,但是测量总体的参数仍然是该总体的一个描述性测量值.

如果一个理论方程被赋予不同参数时具有不同的函数形式,则该方程称为一个**方程族**. 因此,我们可以说,二项概率函数定义了一个非常庞大的**二项概率分布族**. 图 11-2 给出其中的 6 个,用图说明 n 和 p 的变化对分布的影响. 其中,在图 11-2(a)中,p 从 0.25 到 0.50 到 0.75 变化(从上到下),n 恒为常数 5;在图 11-2(b)中,p 从 0.25 到 0.50 到 0.75 变化(从上到下),n 恒为常数 10. 图 11-2(a)的分布值是用$(q+p)^5$ 的展开式(见例 11.6)计算得到,而图 11-2(b)的分布值由下面计算得到

$$(q+p)^{10}=q^{10}+10q^9p+45q^8p^2+120q^7p^3+210q^6p^4+252q^5p^5$$
$$+210q^4p^6+120q^3p^7+45q^2p^8+10qp^9+p^{10}$$

从图 11-2 可以看出二项分布的一些特性. 第一,$p=0.50$ 时,分布是对称的;当 $p<0.50$(正偏,见上册,习题 5.6)和 $p>0.50$(负偏,见上册,习题 5.5)时,分布不对称. 第二,如果 p 不变且 n 增大,则 μ 和 σ 都增大. 第三,如果 n 不变且 p 增大,则 μ 也增大;而 σ 先增大,直到 $p=0.50$,然后对称的减小.

11.10 二项累积概率表

在解决实际问题中,经常需要知道一个二项变量(即 n 次试验成功的次数)小于或等于某个整数的概率,诸如掷一枚硬币 7 次至多出现 2 次正面的概率. 计算这样的概率,要用到离散

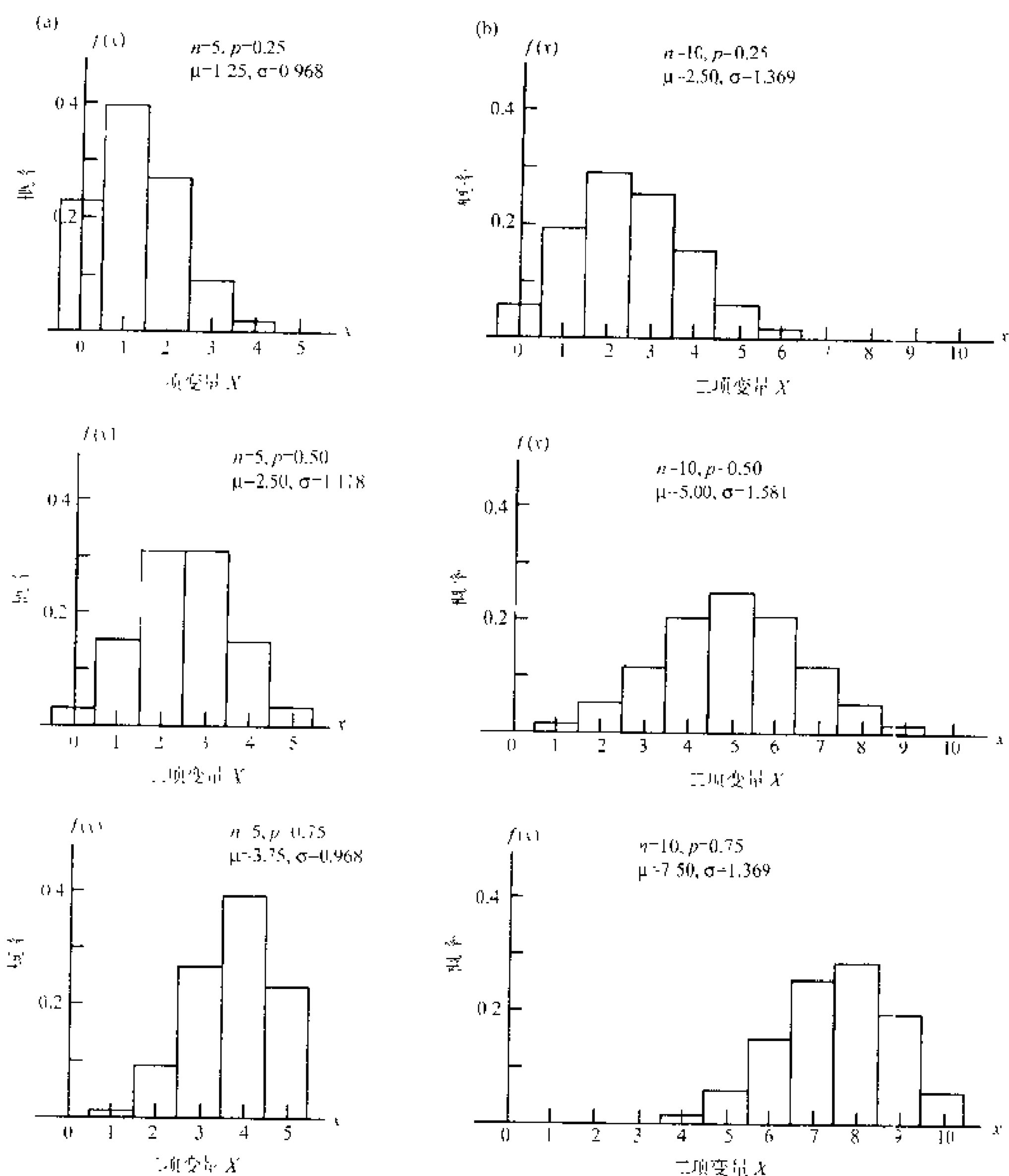


图 11-2

随机变量对任意实数 a 的累积分布函数[上册,公式(10.3)]

$$F(a) = \sum_{x \leq a} f(x)$$

这里,$F(a)$是随机变量取值小于等于 a 的概率,$f(x)$是随机变量取 x 值的概率.用公式(11.2)替换上式中的 $f(x)$,解得二项随机变量的累积分布函数:

$$F(a) = \sum_{x \leq a} \left[f(x) = \binom{n}{x} p^x q^{n-x} \right] \tag{11.10}$$

例 11.8 对二项变量:掷 7 次硬币出现正面的次数,用表 11.2 的二项分布和公式(11.10)计算至多出现 2 次正面的概率.

解 本题需要用到累积分布函数,这里

$$F(2) = P(X \leq 2) = \sum_{x \leq 2} f(x)$$

表 11.2

正面的次数 x	概率 $f(x)=\binom{n}{x}p^xq^{n-x}$
0	$f(0)=\binom{7}{0}(1/2)^0(1/2)^{7-0}=0.00781$
1	$f(1)=\binom{7}{1}(1/2)^1(1/2)^{7-1}=0.05469$
2	$f(2)=\binom{7}{2}(1/2)^2(1/2)^{7-2}=0.16406$
3	$f(3)=\binom{7}{3}(1/2)^3(1/2)^{7-3}=0.27344$
4	$f(4)=\binom{7}{4}(1/2)^4(1/2)^{7-4}=0.27344$
5	$f(5)=\binom{7}{5}(1/2)^5(1/2)^{7-5}=0.16406$
6	$f(6)=\binom{7}{6}(1/2)^6(1/2)^{7-6}=0.05469$
7	$f(7)=\binom{7}{7}(1/2)^7(1/2)^{7-7}=0.00781$
$\sum$	1.00000

因此，代入表 11.2 的数值(近似到 4 位小数)，

$$F(2)=f(0)+f(1)+f(2)$$

$$=0.0078+0.0547+0.1641=0.2266$$

附录表 A.3(**累积二项概率表**)给出由公式(11.10)计算出的累积概率值 $F(a)$. 其中，二项分布有参数值 n(左列)和 p(首行)： $n=2, 3, \cdots, 10$; $p=0.01, 0.05, \cdots, 0.50$.

例 11.9　对二项变量：掷一枚硬币 7 次出现正面的次数，用附录表 A.3 计算至多出现 2 次正面的概率.

解 ☞　对这个问题，我们将表 A.3 要用到的部分复制成表 11.3：给出 $n=7$, $p=0.01, 0.05, \cdots 0.50$ 且 $a=0,1,\cdots,7$ 的累积概率. 从例 11.8 知

$$F(2)=P(X\leqslant 2)=0.2266$$

查表 11.3，从行($a=2$)和列($p=0.50$)的交叉处(右边划圈数字)也得到 $F(2)=0.2266$.

表 11.3

n	a	p 0.01	0.05	0.10	0.15	0.20	0.25	0.30	1/3	0.35	0.40	0.45	0.50
7	0	0.9321	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0585	0.0490	0.0280	0.0152	0.0078
	1	0.9980	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2634	0.2338	0.1586	0.1024	0.0625
	2	1.0000	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.5706	0.5323	0.4199	0.3164	0.2266
	3	1.0000	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.8267	0.8002	0.7102	0.6083	0.5000
	4	1.0000	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9547	0.9444	0.9037	0.8471	0.7734
	5	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9962	0.9931	0.9910	0.9812	0.9643	0.9375
	6	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9995	0.9994	0.9984	0.9963	0.9922
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

为节省篇幅，统计用表往往只给出基本表，而其它可从基本表计算得到. 于是，表 A.3 没有 $p>0.50$ 的累积概率，因为对某个 n 值，对应概率可以从表中给出的概率计算获得. 为求 $p>0.50$ 时的 $F(a)$：查表 A.3 找到合适的 n，然后用 $n-(x+1)$ 替换 a 值，$1-p$ 替换 p 值. 最

后,在行 $n-(x+1)$ 和列 $1-p$ 的交叉处找到累积概率值,并用 1 减去该值即可.

例 11.10 用如表 11.3 所示的表 A.3,对 $n=7$ 和 $p=0.7$ 求 $F(a=3)$

解 对于本题,行 $[7-(3+1)=3]$ 和列 $(1-0.7=0.3)$ 的交叉处(左边划圈数字)是 0.8740,因此

$$F(3)=1-0.8740=0.1260$$

为验证用表 A.3 的解答,对 $n=7$ 和 $p=0.7$ 我们直接用公式(11.10)计算 $F(3)$.

$$\begin{aligned}F(3)&=P(X\leqslant 3)=f(0)+f(1)+f(2)+f(3)\\&=(0.3)^7+7(0.7)(0.3)^6+21(0.7)^2(0.3)^5+35(0.7)^3(0.3)^4\\&=0.00022+0.00357+0.02500+0.09724\\&=0.12603,\text{ 或近似到 4 位有效数字},0.1260\end{aligned}$$

11.11 批验收抽样

批验收抽样(或简称为**验收抽样**)是工业中常使用的一种质量控制方法.它使用二项分布制订**统计决策规则**,据此做出接受或拒绝原材料或产品的决策.质量标准是一个事先规定的集合,如果原材料或产品达到该标准则被接受;否则,被拒绝.

原材料和产品在工业过程中的流动是以某一单位进行的,这一单位就称为**批**.即一定数量的同一产品的集合,诸如 1000 个钢锭,10000 个螺钉,或 5000 个电钻.为确保最终产品的质量,最好的办法就是对每一批的所有元件进行缺陷检验,但这几乎是不可能的.因为,批的数目庞大,这样的检验必定费用昂贵且耗时巨大,而且,某些元件的检验往往会使其报废(例如,炮弹试射检验).因此,对每一批只能进行少量抽样,并且只对这些样本进行缺陷检验.在抽样前,需要制定**抽样计划**以明确样本容量(n)和**接收数**(a).如果样本中有 a 个或更少的不合格品,则接受抽样的整个批;但是,如果样本中有多于 a 个不合格品,则拒绝相应批.

制定工业过程某一阶段上合适的抽样计划,需要使用二项分布给出不同抽样计划的**抽查特征曲线**.对参数为 n 和 a 的抽样计划,这些曲线表明:在不同的不合格品率(p)下,一个批被接受的概率(P_a).对如下抽样计划:$n=10$, $a=0$; $n=10$, $a=1$; $n=50$, $a=0$ 和 $n=50$, $a=1$.图 11-3 给出对应的 4 条抽查特征曲线,其中纵轴为 P_a,横轴为 p.

为加深对这些曲线的理解,我们对 $n=10$ 和 $a=1$ 的抽样计划进行推导和计算.这一抽样计划表明:从某一批随机抽取 10 个个体作为样本,如果样本中有 1 个或没有(0 个)不合格品,则该批被接受(多于 1 个则该批被拒绝).尽管事先不知道批的确切的不合格品率(p),但是如果抽样计划满足多重 Bernoulli 试验(见 11.2 节)的要求,则在给定 p 值下,可以使用二项分布来确定批的接受概率(P_a).通常,这类抽样是无放回的,但在任何时候样本容量 n 总不超过批(总体)的 5%($n\leqslant 0.05N$, 见 11.2 节),因此可以认为 p 在独立试验中基本上保持不变,从而抽样是一个二项试验.如果这样的话,对某一假定的 p 值

$$P_a=P(X\leqslant a)=F(a)=\sum_{x\leqslant a}\left[f(x)=\frac{n!}{x!(n-x)!}p^x q^{n-x}\right]\qquad(11.11)$$

这样,对 $n=10$, $a=1$ 和 $p=0.01$,

$$\begin{aligned}P_a&=\frac{10!}{0!(10-0)!}(0.01)^0(0.99)^{10}+\frac{10!}{1!(10-1)!}(0.01)^1(0.99)^9\\&=(0.99)^{10}+10(0.01)(0.99)^9=0.9957\end{aligned}$$

现在来看图 11-3,你会发现,这就是抽样计划($n=10$, $a=1$)的 P_a 值对 $p=0.01$ 的散点图.而且,对 $n=10$, $p=0.01$,查表 A.3 找到 $F(1)$,证实恰是这个数.

用同样的方法可以找到图 11-3 中所有的散点值.又如,考虑 $p=0.06$ 的抽样计划($n=50$, $a=1$):

$$P_a = \frac{50!}{0!(50-0)!}(0.06)^0(0.94)^{50} + \frac{50!}{1!(50-1)!}(0.06)^1(0.94)^9$$

$$= (0.94)^{50} + 50(0.06)(0.94)^{49} = 0.1900$$

从图 11-3 中曲线，可以看出批验收抽样的两个重要的一般特性. 第一，对给定的样本容量 (n) 和 p 值，接收数 (a) 越大，P_a 值越大. 例如，设 $p=0.04$，样本容量 $n=50$，则 P_a 从 $a=0$ 时的 0.1299 增大到 $a=1$ 时的 0.4005. 第二，对给定的接收数 (a) 和 p 值，样本容量 (n) 越大，P_a 值越小. 例如，设 $p=0.04$，接收数 $a=1$，则 P_a 从 $n=10$ 时的 0.9418 减小到 $n=50$ 时的 0.4005.

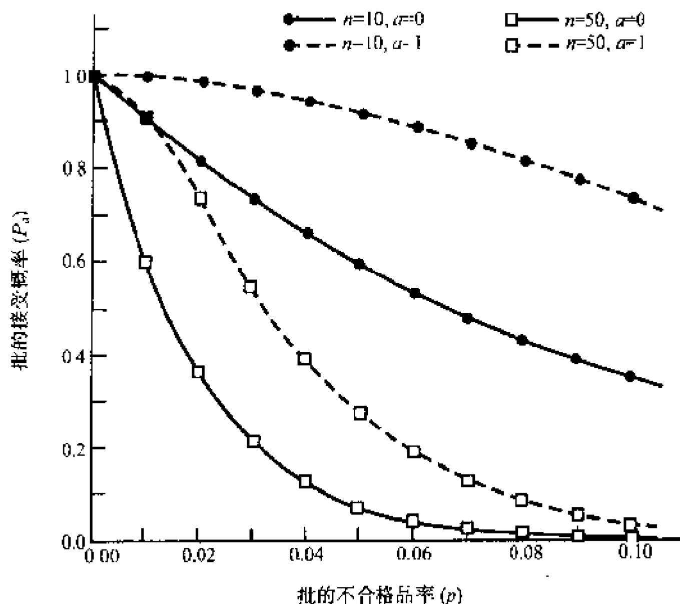


图 11-3

11.12　使用方风险和生产方风险

要在工业过程的某个阶段制定合适的抽样计划，质量控制工程必须考虑许多因素：检验每一样品的成本，工作量，以及所需时间；公司对某一批能够容许的最大不合格品率；供应方能够达到的最小不合格品率；等等. 如果把批的接收者称为使用方，则**使用方风险**是以高过使用方能够容许的 p 值接受一个批的概率. 如果把批的供应者称为生产方，则**生产方风险**是实际上满足使用方要求的一个批被拒绝的概率.

为理解使用方风险，我们假设使用方容许的批不合格品率不超过 $p=0.01$. 这样，如果某一批实际上有 2% 的不合格品，图 11-3 给出 4 个抽样计划的使用方风险：对 $n=10$，$a=0$，使用方风险 $=P_a=0.8171$；对 $n=10$，$a=1$，使用方风险 $=P_a=0.9838$；对 $n=50$，$a=0$，使用方风险 $=P_a=0.3642$；对 $n=50$，$a=1$，使用方风险 $=P_a=0.7358$. 从图 11-3 中曲线可以看出，每个抽样计划的使用方风险随 p 增加而减小.

为理解生产方风险，我们假设生产方确实达到了 $p=0.01$ 的要求. 此时，生产方风险是含有 1% 不合格品的批被拒绝的概率，也是 1 减去批被接受的概率. 因此，对图 11-3 给出的 $p=0.01$ 的批的例子：对 $n=10$，$a=0$，生产方风险 $=1-P_a=1-0.9044=0.0956$；对 $n=10$，$a=1$，生产方风险 $=1-P_a=1-0.9958=0.00042$；对 $n=50$，$a=0$，生产方风险 $=1-P_a=1-0.6050=0.3950$；对 $n=50$，$a=1$，生产方风险 $=1-P_a=1-0.9106=0.0894$；从中可以看出，在使用方要求的某个 p 值下，生产方风险随 n 增加或 a 减少而增加.

例 11.11　假设你是使用方，允许批不合格品率不超过 $p=0.05$，使用的抽样计划是 $n=100$ 和 $a=3$. 如

果某一批的实际不合格品率为 $p=0.06$,使用方风险是多少?

解 ☞ 问题是要计算 $n=100$ 和 $a=3$ 时接受一个不合格品率为6%的批的概率.

$$P_a=\sum_{x\le 3}\left[f(x)=\frac{n!}{x!(n-x)!}p^x q^{n-x}\right]$$

$$=\frac{100!}{0!(100-0)!}(0.06)^0(0.94)^{100}+\frac{100!}{1!(100-1)!}(0.06)^1(0.94)^{99}$$

$$+\frac{100!}{2!(100-2)!}(0.06)^2(0.94)^{98}+\frac{100!}{3!(100-3)!}(0.06)^3(0.94)^{97}$$

$$=0.1430$$

11.13 多元概率分布和联合概率分布

统计方法可以根据分析变量的个数进行分类.当仅有一个分析变量时,统计方法称为**一元统计学**;当有两个分析变量时,称为**二元统计学**;当有多于两个分析变量时,称为**多元统计学**.这些区别只是相对的,在这之前所考虑的离散型概率分布,我们使用的全部是一元方法.现在,对接下来考虑的离散型概率分布——**多项分布**(见 11.14 到 11.19 节)——我们将使用多元方法.

根据所考虑的离散随机变量的个数,离散型概率分布分为一元,二元或多元.这样,二项分布是**离散型一元概率分布**,因为它给出单个离散随机变量 X(成功次数)的所有可能取值的概率[公式(11.2)]

$$f(x)=P(X=x)=\binom{n}{x}p^x q^{n-x}$$

如果试验仅有一个感兴趣的结果,或者尽管有几类结果但希望分别考虑,此时,使用上述的一元概率分布是合适的.然而,更经常的情形,感兴趣的是试验产生的几个不同结果之间的相互关系,此时,如果使用不同的随机变量测量各个结果,则可以计算**离散型二元**或**多元概率分布**,它们给出这些变量的所有可能的组合结果的概率.于是,如果两个结果用离散随机变量 X 和 Y 度量,并且可取确定值 $X=x$ 和 $Y=y$,则 X 和 Y 的离散型概率分布定义为

$$f(x,y)=P(X=x,Y=y)$$

并有性质

(1) $f(x,y)\geqslant 0$,

(2) $\sum_{f(x,y)>0}f(x,y)=1$.

上式表明,所有 $f(x,y)$ 值均大于或等于零,且所有大于零的 $f(x,y)$ 值的和等于 1.

类似地,如果每次试验有 k 个不同结果,分别用离散随机变量 $X_1,X_2,\cdots,X_k$ 度量,且变量可取确定值 $X_1=x_1$, $X_2=x_2$,…, $X_k=x_k$,则这 k 个变量的离散型多元概率分布可定义为

$$f(x_1,x_2,\cdots,x_k)=P(X_1=x_1,X_2=x_2,\cdots,X_k=x_k)$$

并有性质

(1) $f(x_1,x_2,\cdots,x_k)\geqslant 0$,

(2) $\sum_{f(x_1,x_2,\cdots,x_k)>0}f(x_1,x_2,\cdots,x_k)=1$.

上式表明,所有 $f(x_1,x_2,\cdots,x_k)$ 值均大于或等于 0,且取值大于 0 的所有 $f(x_1,x_2,\cdots,x_k)$ 值的和等于 1.

在二元和多元情形,需要同时考虑两个或多个随机变量的相关或**联合结果**.因此,度量且确定二元或多元概率分布的离散随机变量称为**有联合分布的离散随机变量**,这个分布称为离

散型联合概率分布或**离散型联合概率质量函数**.如果从讨论的上下文可明显看出正在处理离散随机变量,也可以简单的将这个分布称为**联合概率分布**或**联合概率函数**.

为理解以上概念,我们来考虑一个具体例子.设碗中放有10粒除颜色外别无差异的石子:4粒为红色(记为R),3粒为绿色(记为G),3粒为蓝色(记为B).试验是闭上眼睛从碗中相继取出5粒石子,每次观测颜色并将取出的石子放回.现在,如果只考虑离散随机变量(X=红色个数),X可取确定值$x=0,1,\cdots,4$,则可以确定离散型一元二项概率分布.如果不这样,而是考虑三个随机变量X_1=红色个数,X_2=绿色个数,X_3=蓝色个数,每次试验它们可取确定值$x=0, 1, \cdots, 4$, $x_2=0, 1, \cdots, 3$和$x_3=0, 1, \cdots, 3$(原文此处有误——译者注),则可以确定多元概率分布

$$f(x_1, x_2, x_3) = P(X_1 = x_1, X_2 = x_2, X_3 = x_3)$$

如果感兴趣的是联合结果(2R, 2G和1B)的概率,由此可解得

$$f(2,2,1) = P(X_1 = 2, X_2 = 2, X_3 = 1)$$

这个分布/函数称为联合概率分布或联合概率函数.对于这个取石子问题,具体的联合概率函数就是11.16节给出的多元概率函数.

11.14　多项试验

二项概率分布基于二项试验(见11.3节),相类似,接下来讨论的离散型概率分布,即**多项分布**,基于**多项试验**.多项试验有以下性质:

(1) 由n次相同的试验组成.

(2) 对每次试验,存在k个互斥且完备的可能结果$A_1, A_2, \cdots, A_k$.

(3) 对每次试验,A_i的概率是p_i(这里$i=1, 2, \cdots, k$)且p_i在各次试验保持不变.

(4) 对每次试验,

$$\sum_i p_i = p_1 + p_2 + \cdots + p_k = 1$$

(5) 各次试验相互独立;任何事件或事件序列不影响随后的事件.

(6) 离散随机变量$X_1, X_2, \cdots, X_k$用于记录n次试验中$A_1, A_2, \cdots, A_k$的次数,实际记录值用$X_1=x_1, X_2=x_2, \cdots, X_k=x_k$表示.这些变量称为**多项随机变量**.

如果将多项试验的性质与二项试验的性质(见11.2和11.3节)相比较,你会明白将**多项试验**称为**二项试验的推广**的原因.两种试验之间的关系的另一种表述是:**二项试验**是$k=2$时的**多项试验的特例**.对于这个二项特例,多项性质可改写为:

(1) 由n次相同的试验组成.

(2) 对每次试验,有$k=2$个可能的互斥事件$A_1=\{成功\}$和$A_2=\{失败\}$.

(3) 对每次试验,A_i的概率是p_i($p_1=p$=成功的概率;$p_2=q$=失败的概率)且p_i在各次试验保持不变.

(4) 对每次试验,

$$\sum_{i=1}^{2} p_i = p - q = 1$$

(5) 各次试验相互独立.

(6) 单个离散随机变量X用于记录n次试验中出现$A_1=\{成功\}$的次数,实际记录值用$X=x$表示.

11.15　多项系数

多项试验是二项试验(见11.14节)的推广,同样,**多项系数**也是二项系数[公式11.1]的

推广.多项系数记录 n 个个体的可能排列数,这里 n 个个体由 k 类构成:类型 1 有 x_1 个,类型 2 有 x_2 个,…,类型 k 有 x_k 个.如果 x_i 是非负整数且 $x_1+x_2+\cdots+x_k=n$,则多项系数定义为

$$\text{多项系数}=\binom{n}{x_1,\ x_2,\ \cdots,\ x_k}=\frac{n!}{x_1!x_2!\cdots x_k!} \tag{11.12}$$

例 11.12 如 11.13 节,一只碗中放有 10 粒石子(4R,3G,3B).从碗中相继取出 10 粒,共有多少种方法(即有多少种方法使 3 种颜色的 10 粒石子按次序排成一列)?

解 问题是:4 粒红色石子(记为 R),3 粒绿色石子(记为 G)和 3 粒蓝色石子(记为 B)有多少种排列?如果定义:$X_1=R$ 的个数 $=x_1=4$,$X_2=G$ 的个数 $=x_2=3$,$X_3=B$ 的个数 $=x_3=3$,由多项系数求得排列数

$$\binom{n}{x_1,\ x_2,\ x_3}=\frac{n!}{x_1!x_2!x_3!}$$

$$\binom{10}{4,\ 3,\ 3}=\frac{10!}{4!3!3!}=\frac{3628800}{864}=4200$$

11.16 多项概率函数

由前知,多项试验是二项试验(见 11.14 节)的推广,多项系数是二项系数(见 11.15 节)的推广,因此,**多项概率函数**也是二项概率函数[公式(11.2)]的推广.二项概率函数

$$f(x)=\binom{n}{x}p^xq^{n-x},\ x=0,\ 1,\ 2,\ \cdots,\ n$$

给出在二项试验的 n 次试验中出现 x 次成功的概率,而多项概率函数给出在多项试验的 n 次试验中观测到事件 A_1 发生 x_1 次,A_2 发生 x_2 次,…,A_k 发生 x_k 次的概率.这个联合概率函数(见 11.13 节)是

$$f(x_1,\ x_2,\ \cdots,\ x_k)=\binom{n}{x_1,\ x_2,\ \cdots,\ x_k}p_1^{x_1}p_2^{x_2}\cdots p_k^{x_k}=\frac{n!}{x_1!x_2!\cdots x_k!}p_1^{x_1}p_2^{x_2}\cdots p_k^{x_k} \tag{11.13}$$

这里 x_i 是 n 次试验中事件 A_i 发生的次数,p_i 是每次试验中 A_i 发生的概率.$x_1+x_2+\cdots+x_k=n$,且 $p_1+p_2+\cdots+p_k=1$.

例 11.13 同例 11.12,一只碗中放有 10 粒石子(4 粒红色(记为 R),3 粒绿色(记为 G)和 3 粒蓝色(记为 B)).试验是从碗中相继取出 5 粒石子,每次观测颜色并放回.这个试验满足多项性质:(1)是 $n=5$ 重试验;(2)对每次试验,有 3 个互斥且完备的可能结果:$A_1=\{R\}$,$A_2=\{G\}$,$A_3=\{B\}$;(3)对每次试验,$P(A_1)=p_1=4/10$,$P(A_2)=p_2=3/10$,$P(A_3)=p_3=3/10$;(4)对每次试验,$\sum p_i=4/10+3/10+3/10=1$;(5)各次试验相互独立;(6)离散随机变量:$X_1=R$ 的个数,$X_2=G$ 的个数和 $X_3=B$ 的个数,用于记录 n 次试验中出现 A_1,A_2 和 A_3 的次数.试问:在 5 次选取中,得到 $X_1=x_1=2$,$X_2=x_2=2$ 和 $X_3=x_3=1$ 的概率是多少?分别使用:(a)概率原理和(b)多项概率函数[公式(11.13)]计算这个概率.

解 (a) 5 个事件的序列之一是 RRGGB.对每次试验,已知 p_i 值且各次试验相互独立,由特殊乘法原理的推广[上册,公式(9.9)],得到

$$P(R\cap R\cap G\cap G\cap B)=P(R)P(R)P(G)P(G)P(B)$$
$$=\left(\frac{4}{10}\right)\left(\frac{4}{10}\right)\left(\frac{3}{10}\right)\left(\frac{3}{10}\right)\left(\frac{3}{10}\right)=0.00432$$

可以写成多项形式

$$P(A_1 \cap A_1 \cap A_2 \cap A_2 \cap A_3) = p_1^2 p_2^2 p_3^1$$
$$= \left(\frac{4}{10}\right)^2 \left(\frac{3}{10}\right)^2 \left(\frac{3}{10}\right) = 0.00432$$

但是,这仅仅是产生 $X_1=2,\ X_2=2,\ X_3=1$ 结果的5个事件的多种可能序列之一.精确个数要由多项系数计算[公式(11.12)].即

$$\binom{n}{x_1,\ x_2,\ x_3} = \frac{n!}{x_1!x_2!x_3!} = \frac{5!}{2!2!1!} = 30$$

合并这些结果,得到

$$P(X_1 = 2,\ X_2 = 2,\ X_3 = 1) = (30)(0.00432) = 0.12960$$

(b) 直接用公式(11.13),得到同样结果:

$$f(x_1,\ x_2,\ x_3) = \binom{n}{x_1,\ x_2,\ x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3} = \frac{n!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

$$f(2,\ 2,\ 1) = \binom{5}{2,\ 2,\ 1}\left(\frac{4}{10}\right)^2\left(\frac{3}{10}\right)^2\left(\frac{3}{10}\right)^1 = \frac{5!}{2!2!1!}\left(\frac{4}{10}\right)^2\left(\frac{3}{10}\right)^2\left(\frac{3}{10}\right)^1$$
$$= (30)(0.00432) = 0.12960$$

11.17　多项概率分布族

回顾11.9节,在二项概率函数[公式(11.2)]中,n 和 p 是参数且为确定常数,并且 n 和 p 的每个组合就定义了二项分布族的一员.类似地,$n,\ p_1,\ p_2,\ \cdots,\ p_k$ 称为多项概率函数[公式(11.13)]的参数,并且 $n,\ p_1,\ p_2,\ \cdots,\ p_k$ 值的每个组合也定义了有无穷多个的**多项分布族**的一员.每个多项分布都可以表示为指定参数的具体函数,或者表示成完整的概率分布列或概率表.

例11.14　同例11.13,一只碗中放有10粒石子(4粒红色(记为R),3粒绿色(记为G)和3粒蓝色(记为B)).试验是从碗中相继取出2粒,每次观测石子颜色并放回.于是 $n=2$;$A_1=\{R\}$,$A_2=\{G\}$,$A_3=\{B\}$;$X_1=$R的个数$=x_1$,$X_2=$G的个数$=x_2$,和 $X_3=$B的个数$=x_3$;$p_1=0.4$,$p_2=0.3$ 和 $p_3=0.3$.试用一个概率表给出这个试验的多项分布.

解　所求的多项分布在表11.4给出.

表11.4

R,G,B的个数 $x_1,\ x_2,\ x_3$	概率 $f(x_1,\ x_2,\ x_3)=\binom{n}{x_1,\ x_2,\ x_3}p_1^{x_1}p_2^{x_2}p_3^{x_3}$
1,1,0	$f(1,\ 1,\ 0)=\binom{2}{1,\ 1,\ 0}(0.4)^1(0.3)^1(0.3)^0=(2)(0.4)(0.3)(1)=0.24$
1,0,1	$f(1,\ 0,\ 1)=\binom{2}{1,\ 0,\ 1}(0.4)^1(0.3)^0(0.3)^1=(2)(0.4)(1)(0.3)=0.24$
0,1,1	$f(0,\ 1,\ 1)=\binom{2}{0,\ 1,\ 1}(0.4)^0(0.3)^1(0.3)^1=(2)(1)(0.3)(0.3)=0.18$
2,0,0	$f(2,\ 0,\ 0)=\binom{2}{2,\ 0,\ 0}(0.4)^2(0.3)^0(0.3)^0=(1)(0.16)(1)(1)=0.16$
0,2,0	$f(0,\ 2,\ 0)=\binom{2}{0,\ 2,\ 0}(0.4)^0(0.3)^2(0.3)^0=(1)(1)(0.09)(1)=0.09$
0,0,2	$f(0,\ 0,\ 2)=\binom{2}{0,\ 0,\ 2}(0.4)^0(0.3)^0(0.3)^2=(1)(1)(1)(0.09)=0.09$
$\sum$	1.00

11.18　多项概率分布的均值

多项概率分布对其每一个离散随机变量都有均值(期望值),这个均值对应于随机变量在 n 次试验中出现的期望次数.设分布的随机变量是 $X_1, X_2, \cdots, X_k$,则**多项概率分布的均值**是

$$E(X_1)=\mu_1=np_1,\ E(X_2)=\mu_2=np_2,\ \cdots,\ E(X_k)=\mu_k=np_k \qquad (11.14)$$

例 11.15　对表 11.4 中的多项分布,X_1、X_2 和 X_3 的均值分别是多少?

解 ☞　均值分别是

$$E(X_1)=\mu_1=np_1=2(0.4)=0.8$$

$$E(X_2)=\mu_2=np_2=2(0.3)=0.6$$

$$E(X_3)=\mu_3=np_3=2(0.3)=0.6$$

11.19　多项式展开和多项式定理

多项概率分布与**多项式展开**和**多项式定理**有关,就像二项概率分布与二项式展开和二项式定理(见 11.7 节)有关一样.

形如 $a_1+a_2+\cdots+a_k$ 的**多项代数表达式**(见上册,1.11 节)自乘 n 次之后产生的若干项和,称为多项式展开.如果 $a_1, a_2, \cdots, a_k$ 是实数且 n 是正整数,展开 $(a_1+a_2+\cdots+a_k)^n$ 的一般公式

$$(a_1+a_2+\cdots+a_k)^n=\sum_{x_1+x_2+\cdots+x_k=n}\binom{n}{x_1,\ x_2,\ \cdots,\ x_k}a_1^{x_1}a_2^{x_2}\cdots a_k^{x_k} \qquad (11.15)$$

称为多项式定理.这里,符号 $\sum_{x_1+x_2+\cdots+x_k=n}$ 表示:对和为 n 的 k 个非负整数的所有可能组合进行求和.

例 11.16　对表 11.4 中的多项分布,令 $n=2$,$a_1=p_1$,$a_2=p_2$ 和 $a_3=p_3$,试说明多项概率分布与多项展开式和多项式定理之间的关系.

解 ☞　将上述参数代入公式(11.15),对满足 $x_1+x_2+x_3=2$ 的非负整数 x_1,x_2 和 x_3 的所有可能组合,多项展开式为

$$(p_1+p_2+p_3)^2=\sum_{x_1+x_2+x_3=2}\binom{2}{x_1,\ x_2,\ x_3}p_1^{x_1}p_2^{x_2}p_3^{x_3}$$

$$=\binom{2}{1,\ 1,\ 0}p_1^1p_2^1p_3^0+\binom{2}{1,\ 0,\ 1}p_1^1p_2^0p_3^1+\binom{2}{0,\ 1,\ 1}p_1^0p_2^1p_3^1$$

$$+\binom{2}{2,\ 0,\ 0}p_1^2p_2^0p_3^0+\binom{2}{0,\ 2,\ 0}p_1^0p_2^2p_3^0+\binom{2}{0,\ 0,\ 2}p_1^0p_2^0p_3^2$$

$$=2p_1p_2+2p_1p_3+2p_2p_3+p_1^2+p_2^2+p_3^2$$

从上式可以发现,对于表 11.4 中多项分布的每一个概率值,在多项展开式中都有对应的一项.事实上,多项概率分布的名称正是缘于它与多项式展开的这种关系.

11.20　超几何试验

同本章前面的离散型概率分布(见 11.3 和 11.4 节)一样,接下来考虑的**超几何概率分布**

也是基于一个试验:**超几何试验**.超几何试验类似于二项试验,只是涉及的是**无放回**的有限总体,各次试验相互不独立,从而"成功"的概率随着样本从总体中移走而改变.超几何试验的特征是:

(1) 包含 n 个个体的随机样本是从包含 N_T 个个体的有限总体中无放回的一次抽取一个(n 次试验)得到.

(2) 总体的 N_T 个个体中,N_S 是一类,称为"成功";N_F 是另一类,称为"失败"($N_S+N_F=N_T$).

(3) 离散随机变量 X,这里称为**超几何随机变量**,用于记录样本中成功的次数($X=x$).

11.21 超几何概率函数

超几何概率函数用于确定在超几何试验的 n 次试验中出现 x 次成功的概率.超几何概率函数不同于二项概率函数[公式(11.2)],因为它采取无放回抽样,每一次试验后成功的概率都在改变.因此,超几何概率函数使用一种不同的计算方法,基于三个概率概念:古典概率比[上册,公式(8.1)],计数规则:乘法原理(见上册,9.12 节),和计数规则:组合(见上册,9.14 节).

在概率的古典解释(见上册,8.1 节)中,事件 A 的概率[记作 $P(A)$]是有利于 A 的可能结果的个数(用 N_A 表示)对试验的所有可能结果的总数(用 N 表示)的比,并且假设所有结果等可能地出现.即

$$P(A)=\frac{N_A}{N}$$

这个比例是超几何概率函数的基本结构,用 N_A 表示在超几何试验的 n 次试验中出现 x 次成功的方法数,N 表示在 n 次试验中出现的结果总数.计算 N_A 和 N 需要用到上述两个计数规则.

为理解这一过程,再次考虑例 11.4 的试验:从放有 20 粒石子(12R,8G)的碗中相继取出 6 粒,计算在 6 次选取中取出 4R 的概率.在例 11.4,抽样是有放回的,在各次试验中概率保持不变,从而使用二项概率函数计算出 $f(4)=0.311$.现在,假设无放回抽样,考虑 6 次选取中取出 4R 的概率.此时,试验满足超几何试验(见 11.20 节)的条件,其中:$n=$ 取石子次数 $=6$,$N_T=$ 石子总数 $=20$,$N_S=$ R 的个数 $=12$,$N_F=$ G 的个数 $=8$,$x=$ 成功的次数 $=4$.由此,计算比例

$$P(\text{取出 4R})=P(X=4)=\frac{N_A}{N}$$

这里,N_A 是在 6 次选取中取出 4R 的总方法数,N 是在 6 次选取中能够得到的可能结果的总数.为得到 N_A,我们使用两个计数规则.首先,使用计数规则:组合,计算在 6 次选取中能从 12R 取出 4R 的方法数,

$${}_{N_S}C_x={}_{12}C_4=\frac{N_S!}{x!(N_S-x)!}=\frac{12!}{4!(12-4)!}=\frac{11880}{24}=495$$

以及在 6 次选取中能从 8G 取出 2G 的方法数,

$${}_{N_F}C_{n-x}={}_8C_2=\frac{N_F!}{(n-x)![N_F-(n-x)]!}=\frac{8!}{2!(8-2)!}=\frac{56}{2}=28$$

再使用计数规则:乘法原理计算 N_A:

$$N_A = {}_{N_S}C_x \times {}_{N_F}C_{n-x} = 495 \times 28 = 13860$$

为得到 N,即在 6 次选取中出现的可能结果总数,再次使用计数规则:组合,即

$$N = {}_{N_T}C_n = {}_{20}C_6 = \frac{N_T!}{n!(N_T - n)!} = \frac{20!}{6!(20-6)!} = \frac{27907200}{720} = 38760$$

因此,
$$P(X = 4) = \frac{13860}{38760} = 0.3576$$

对这个例子进行总结,如果超几何分布要求的条件能够满足,并且用一个概率函数对超几何变量定义的样本空间中的每一点指定概率值,则这个函数称为**超几何概率分布**或**超几何分布**或**超几何概率质量函数**或简称为**超几何概率函数**.有如下定义:

$$f(x) = P(X = x) = \frac{{}_{N_S}C_x \times {}_{N_F}C_{n-x}}{{}_{N_T}C_n} = \frac{\binom{N_S}{x}\binom{N_F}{n-x}}{\binom{N_T}{n}} \tag{11.16}$$

由上册公式(9.20),得到

$$f(x) = P(X = x) = \frac{\left[\frac{N_S!}{x!(N_S - x)!}\right]\left[\frac{N_F!}{(n-x)![N_F - (n-x)]!}\right]}{\left[\frac{N_T!}{n!(N_T - n)!}\right]},\ x = 0, 1, 2, \cdots, n \tag{11.17}$$

这里:$x \leqslant N_S$(样本中成功的次数必然小于或等于总体中成功的次数);$n - x \leqslant N_F$(样本中失败的次数必然小于或等于总体中失败的次数);N_T, n 和 N_S 是正整数,且满足 $N_S \leqslant N_T$ 和 $n \leqslant N_T$.

例 11.17　某经理需要从 6 个男性员工(记为 M)和 4 个女性员工(记为 W)中选出一个 3 人委员会.他将员工名字分别写在同样的 6 张纸条上,并将纸条放入一只碗中,然后闭上眼睛从中取出 3 张.分别使用两种方法:(a) 上册第九章的条件概率方法和(b)超几何概率函数,计算取出 2 个女性的概率.

解 ☞　(a) 包含 2 个女性的选择序列有 3 个:$W \cap W \cap M$, $W \cap M \cap W$ 和 $M \cap W \cap W$. 由上册公式(9.5),

$$P(W \cap W \cap M) = \frac{4}{10} \times \frac{3}{9} \times \frac{6}{8} = 0.10$$

$$P(W \cap M \cap W) = \frac{4}{10} \times \frac{6}{9} \times \frac{3}{8} = 0.10$$

$$P(M \cap W \cap W) = \frac{6}{10} \times \frac{4}{9} \times \frac{3}{8} = 0.10$$

然后,由特殊加法原理的推广(上册,8.6 节,性质 4)

$$\begin{aligned} P(2\text{个女性}) &= P(W \cap W \cap M) \cup P(W \cap M \cap W) \cup P(M \cap W \cap W) \\ &= P(W \cap W \cap M) + P(W \cap M \cap W) + P(M \cap W \cap W) \\ &= 0.10 + 0.10 + 0.10 = 0.30 \end{aligned}$$

(b) 因为满足超几何试验的条件,所以可使用公式(11.16).即

$$_{N_T}C_n = {}_{10}C_3 = \frac{10!}{3!(10-3)!} = 120$$

$$_{N_S}C_x \times _{N_F}C_{n-x} = {}_4C_2 \times {}_6C_1 = \left(\frac{4!}{2!2!}\right)\left(\frac{6!}{1!5!}\right) = (6)(6) = 36$$

因此,

$$P(\text{取出 2 个女性}) = \frac{36}{120} = 0.30$$

11.22 超几何概率分布族

同二项(见 11.9 节)和多项(见 11.17 节)分布一样,超几何分布也是一个概率分布族,有参数 N_T, N_S 和 n.

例 11.18 对例 11.17 中试验,将给定的参数值代入公式(11.17),计算超几何概率分布.

解 所求的超几何分布在表 11.5 给出.

表 11.5

女性人数 x	概率 $f(x)=\frac{\binom{N_S}{x}\binom{N_F}{n-x}}{\binom{N_T}{n}}$
0	$f(0)=\frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}}=\frac{(1)(20)}{(120)}=0.16667$
1	$f(1)=\frac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}}=\frac{(4)(15)}{(120)}=0.50000$
2	$f(2)=\frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}}=\frac{(6)(6)}{(120)}=0.30000$
3	$f(3)=\frac{\binom{4}{3}\binom{6}{0}}{\binom{10}{3}}=\frac{(4)(1)}{(120)}=0.03333$
$\sum$	1.00000

11.23 超几何概率分布的均值,方差和标准差

我们不做推导,直接给出超几何分布的均值(期望)

$$E(X) = \mu = \frac{nN_S}{N_T} \tag{11.18}$$

方差

$$\sigma^2 = \frac{nN_SN_F(N_T-n)}{N_T^2(N_T-1)} \tag{11.19}$$

标准差

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{nN_SN_F(N_T - n)}{N_T^2(N_T - 1)}} \tag{11.20}$$

例 11.19 计算表 11.5 所示的超几何分布的均值和方差.

解 由公式(11.18),

$$\mu = \frac{nN_S}{N_T} = \frac{3 \times 4}{10} = 1.2$$

由公式(11.19),

$$\sigma^2 = \frac{nN_SN_F(N_T - n)}{N_T^2(N_T - 1)} = \frac{3 \times 4 \times 6 \times (10 - 3)}{10^2(10 - 1)} = 0.56$$

11.24 超几何概率分布的推广

就像多项概率函数是二项概率函数(见 11.16 节)的推广一样,这里也存在一个类似的联合概率函数(见 11.13 节),即**广义超几何概率函数**. 适用该函数的超几何试验具有以下性质:

(1) 从一个含有 N_T 个个体的总体中无放回抽取含有 n 个个体的随机样本.

(2) 总体的 N_T 个个体:类型 1 有 N_1 个,类型 2 有 N_2 个,…,类型 k 有 N_k 个,并且 $N_1 + N_2 + \cdots + N_k = N_T$.

(3) 离散随机变量 $X_1, X_2, \cdots, X_k$ 用于记录样本中类型 1,2,…, k 的出现个数,实际记录值用 $X_1 = x_1, X_2 = x_2, \cdots, X_k = x_k$ 表示.

联合概率函数给出含有 x_1 个类型 1,x_2 个类型 2,…, x_k 个类型 k 的概率,也简称为超几何概率函数,是

$$f(x_1, x_2, \cdots, x_k) = P(X_1 = x_1, X_2 = x_2, \cdots, X_k = x_k)$$
$$= \frac{\binom{N_1}{x_1}\binom{N_2}{x_2}\cdots\binom{N_k}{x_k}}{\binom{N_T}{n}} \tag{11.21}$$

其中 $x_1 + x_2 + \cdots + x_k = n$.

例 11.20 随机地从一副 52 张扑克牌中无放回地抽取 10 张. 试问:抽出 2 张梅花,2 张红心,2 张方块和 4 张黑桃的概率是多少?

解 我们使用超几何分布求解本问题,这里:X_1 = 样本中梅花数 = x_1 = 2, X_2 = 样本中红心数 = x_2 = 2, X_3 = 样本中方块数 = x_3 = 2, X_4 = 样本中黑桃数 = x_4 = 4, N_T = 52, $N_1 = N_2 = N_3 = N_4$ = 每一花色的张数 = 13 和 n = 10. 将这些值代入公式(11.21),得到

$$f(2, 2, 2, 4) = \frac{\binom{13}{2}\binom{13}{2}\binom{13}{2}\binom{13}{4}}{\binom{52}{10}} = \frac{(78)(78)(78)(715)}{(15820024220)} = 0.0214$$

11.25 超几何分布的二项和多项近似

解决**从有限总体无放回抽样**的两结果或多结果的概率问题,超几何分布是一个可以使用的合适的离散型概率分布. 然而,对很大的总体,有放回抽样和无放回抽样基本上没什么差别. 如果 $n \leqslant 0.05N$,二项分布可用于解决无放回抽样的两结果问题,而多项分布可用于解决无放回抽样的多结果问题. 在二项情形,这种近似称为**超几何分布的二项近似**,在多项情形,称为**超**

几何分布的多项近似.

11.26　Poisson 过程及其随机变量和试验

本章的前三个离散型概率分布处理的是:在固定的试验次数 n 下,随机事件发生的概率.现在,本章的第 4 个也就是最后一个概率分布,**Poisson 概率分布**,将考虑**在连续时间和空间单位**(见上册,2.6 节)**上发生的随机事件**的概率.Poisson 分布与二项分布相似,也是处理两个可能结果之一(仍称为"成功"结果)发生的次数.但是,对于二项分布,成功发生的次数是有限次;对于 Poisson 分布,成功可以是连续时间或空间单位上的无限多次.现实世界中,能够用 Poisson 方法恰当地计算其概率结果的一些离散随机变量的例子有:在上午 10:00 到 10:30 之间到达交换台的电话呼叫次数,置于一种放射物附近的 Geiger 计数器在 10 秒内记录的粒子个数,4 米长的电缆上出现的缺陷个数,1 立方厘米血液中白细胞的个数,一辆新轿车表面的缺陷个数,以及一株紫菜上生长的细菌群体数.

使用 Poisson 方法计算随机变量的概率结果,要求产生这些随机结果的**过程**必须是一个 **Poisson 过程**.一个过程是 Poisson 过程,只须满足以下假设条件:

(1) 在给定的连续时间或空间单位内,已知由经验确定的一个正常数,记为 λ(小写希腊字母 lambda),它是给定单位内**成功的平均出现率**.比例 λ 刻画了观测到的产生成功的过程,而且对类似的确定单位 λ 相等.

(2) 在给定单位的任何大小的子单位内,成功次数与任何其它不重叠子单位内的成功次数**相互独立**.

(3) 如果某个确定单位被分成一些很小的子单位,记为 h,则在 h 内恰好出现一次成功的概率非常小;并且,在该单位的任何 h 内这个概率都相等(为常数),而无论成功在何时(或何位置)出现.随着 h 愈来愈小,在 h 内恰好出现一次成功的微小概率愈来愈接近值 λh.

(4) 在很小的子单位 h 内出现多于一次成功的概率基本上为零.随着 h 越来越小,这一概率也越来越接近于零.

如果在某一时间或空间单位内产生随机事件的过程是 Poisson 过程,并且用一个离散随机变量 X 记录给定单位(或它的子单位)内的成功次数($X=x$),则变量 X 称为 **Poisson 随机变量**,记录成功次数的试验称为 **Poisson 试验**.这些概念以及 **Poisson 概率函数**(见 11.27 节)是以它们的发现者,法国数学家 Siméon D. Poisson(1781—1840)的名字命名的.

为理解上述概念,考虑以下 Poisson 试验.一个电缆制造商,已经知道缺陷是在电缆生产过程中"随机"产生的,想用 Poisson 方法计算一段固定长度(单位)的电缆上出现不同缺陷(成功)个数的概率.他决定用 4 米作为固定单位,在对许多的 4 米长电缆记录缺陷之后,发现这些记录的平均值(算术平均值)是(每 4 米 4.0 个缺陷).以 4 米为一单位,存在一个由经验得到的正常数($\lambda=$ 每 4 米 4.0 个缺陷),λ 是该单位内缺陷(成功)随机出现的平均率,而且,可合理地认为对任何 4 米均相同,因此,假设(1)能够满足.考虑 4 米单位上的一个 10 厘米部分,如果在这个子单位内出现的缺陷个数不依赖于任何其它不重叠的子单位内出现的缺陷个数,则假设(2)能够满足.对 4 米单位上的一个非常小的子单位 h,如果在任何 h 内恰好有一个缺陷的概率 $[P(X=1)\approx\lambda h=(4.0)h]$ 非常小并为常数,而不论 h 在 4 米单位上的位置如何,则假设(3)能够满足.最后,如果在任何子单位 h 内多于一个缺陷的概率基本上为零 $[P(X>1)\approx 0]$,则假设(4)能够满足.

例 11.21　在每个周末,市区一家诊所的接待员记录下本周内因同等强度的高传染疾病来就诊的新病例的个数.在过去三周,这些记录是增加的,分别为 2,10 和 30 例.试解释它不是 Poisson 过程的原因.

解 ☞　这个试验要成为一个 Poisson 试验,在给定时间单位(一周)内,产生成功(就诊的新病例)的过程必须是一个 Poisson 过程,即不能严重违背 4 个 Poisson 假设的任何一个.然而,在本例中至少有三个假设遭到严重违背.

假设(1)实际上是:在给定单位(一周)内,存在一个由经验确定的成功(新病例)出现的平均率 λ,而且它

在所有类似的确定单位内相同. 显然,这里并不是这种情形,因为产生过程明显不稳定并且变化很快. 在这种情形下,对这三周(每周 $\lambda = 42/3 = 14$ 个病例)计算得到的 λ 不能认定在随后几周仍然有效.

假设(2)实际上是:在一周的任何子单位(如周五)内就诊的新病例个数与任何其它非重叠子单位(如周一)的新病例个数之间相互独立. 但是,在某个拥挤的城市内疾病大范围传染的条件下,这个假设很可能被违背. 因为,如果病人在周一受到某种已知疾病感染,他回到居住区将会传染给邻居,结果造成诊所在下周出现新病例.

假设(3)实际上是:在单位(一周)的任何相同且极小的子单位(即以 n 秒记)内恰好出现一次成功的概率非常小且为常数. 但是,在这样的传染条件下,期望这个瞬间概率保持不变是不合理的.

最后,对于假设(4),即病例数可以非常小使得在很小的子单位内有多于一个新病例的概率接近于零. 但是,如果病例数增长很快这也将被违背.

11.27　Poisson 概率函数

11.26 节指出 Poisson 过程的一个假设:在任何给定的时间或空间单位内,由经验可确定成功出现的平均率 λ,并在任何类似定义的单位内 λ 相等. 由这个和另外的三个假设,不做证明,我们给出以下关系:在给定单位的任意倍数(用 t 表示)之内,成功出现的平均(期望)率为 λt. 于是,对 11.26 节的电缆缺陷试验,其中 λ = 每 4 米 4.0 个缺陷,在 24 米($t = 6$ 个单位)内出现缺陷的平均率为

$$\lambda t = (\text{每 4 米 4.0 个缺陷}) \times (6\text{ 单位})$$
$$= \text{每 24 米 24.0 个缺陷}$$

在 1 米$\left(t = \frac{1}{4}\text{ 单位}\right)$内

$$\lambda t = (\text{每 4 米 4.0 个缺陷}) \times \left(\frac{1}{4}\text{ 单位}\right)$$
$$= \text{每米 1.0 个缺陷}$$

Poisson 概率函数用常数 λt 计算在 Poisson 试验中给定单位的倍数 t 内出现($X = x$)次成功的概率. 不经推导,我们给出该函数

$$f(x) = P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!},\ x = 0, 1, 2, \cdots \tag{11.22}$$

这里 e 是自然对数的底,等于 2.71828…[见上册,习题 1.23(b)], 且 $x = 0, 1, 2, \cdots$,表明函数的定义域(见上册,1.17 节)是一切非负整数.

例 11.22　假设 11.26 节描述的电缆缺陷试验是一个 Poisson 试验,则一根 1 米长电缆上出现 2 个缺陷的概率是多少?

解　由公式(11.22),1 米长电缆$\left(t = \frac{1}{4}\text{ 单位}\right)$出现 2 个缺陷的概率是

$$f(2) = P(X = 2) = \frac{\left(4.0 \times \frac{1}{4}\right)^2 e^{-(4.0 \times 1/4)}}{2} = \frac{(1)^2 e^{-1}}{2} = \frac{e^{-1}}{2}$$

利用计算器的 e^x 键,可以直接得到解:

$$f(2) = 0.1839$$

或者,用 $e = 2.71828\cdots$ 和 $e^{-x} = 1/e^x$[见上册,例 1.16(a)],可以得到解:

$$f(2) = \frac{e^{-1}}{2} = \frac{1}{2e} = \frac{1}{2(2.71828)} = 0.1839$$

11.28　Poisson 概率分布族

同本章先前讨论的分布一样,Poisson 分布也是一个概率函数族,只有一个参数 λt.

表 11.6

不合格品数 x	概率 $f(x)=\frac{(\lambda t)^x e^{-\lambda t}}{x!}$
0	$f(0)=\frac{(1.0)^0 e^{-1.0}}{0!}=0.36788$
1	$f(1)=\frac{(1.0)^1 e^{-1.0}}{1!}=0.36788$
2	$f(2)=\frac{(1.0)^2 e^{-1.0}}{2!}=0.18394$
3	$f(3)=\frac{(1.0)^3 e^{-1.0}}{3!}=0.06131$
4	$f(4)=\frac{(1.0)^4 e^{-1.0}}{4!}=0.01533$
5	$f(5)=\frac{(1.0)^5 e^{-1.0}}{5!}=0.00307$
6	$f(6)=\frac{(1.0)^6 e^{-1.0}}{6!}=0.00051$
$\sum$	0.99992

例 11.23 由公式(11.22)确定 11.26 节中电缆缺陷试验的 Poisson 概率分布,其中 λ = 每 4 米 4.0 个缺陷,$t=\frac{1}{4}$ 单位,从而 λt = 每米 1.0 个缺陷.

解 ☞ 因为公式(11.22)的定义域是一切非负整数,即可数多个数值(见上册,10.2 节),所以确定出完整的 Poisson 概率分布是不可能的.由表 11.6 给出 $x=0, 1, \cdots, 6$ 的分布值,可以看出,随着$(X=x)$增加 $P(X=x)$迅速趋于0;也可以看出,取值 $x=0, 1, \cdots, 6$ 的分布中,总概率是多少

$$\sum_{x=0}^{6} f(x) = 0.99992$$

11.29 Poisson 概率分布的均值,方差和标准差

对于 Poisson 分布,λ 是给定单位内成功出现的平均率,λt 是在 t 个单位间隔内成功出现的平均率,于是,不做推导,我们给出

分布的均值 = 在 t 个单位间隔内成功的期望次数

$$\mu = E(X) = \lambda t \tag{11.23}$$

$$\text{分布的方差} = \sigma^2 = \mu = \lambda t \tag{11.24}$$

$$\text{分布的标准差} = \sigma = \sqrt{\sigma^2} = \sqrt{\lambda t} \tag{11.25}$$

11.28 节指出,Poisson 分布只有一个参数 λt. 现在,可以看到,λt 既是分布的均值又是分布的方差.因此,公式(11.22)经常写成

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}, \quad x = 0, 1, 2, \cdots \tag{11.26}$$

例 11.24 表 11.6 给出的 Poisson 分布的均值,方差,标准差和参数各是多少?

解 ☞ 对这个 Poisson 分布,λ 等于 4.0,t 等于$\frac{1}{4}$,均值是

$$\mu = \lambda t = 4.0 \times \frac{1}{4} = 1.0$$

方差是

$$\sigma^2 = \mu = \lambda t = 1.0$$

标准差是

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.0} = 1.0$$

参数是

$$\lambda t = 1.0$$

11.30 Poisson 累积概率表

11.10 节介绍了附录表 A.3(**二项累积概率表**).对某些二项分布,表 A.3 给出由公式(11.10)

$$F(a) = \sum_{x \leq a}\left[f(x) = \binom{n}{x}p^x q^{n-x}\right]$$

计算出的累积概率值.类似地,附录表 A.4(**Poisson 累积概率表**)给出在 $\mu=0.001,\cdots,1.00$ 和 $\mu=1.1,\cdots,8.0$ 时由公式

$$F(a) = \sum_{x \leq a}\left[f(x) = \frac{\mu^x e^{-\mu}}{x!}\right] \tag{11.27}$$

计算出的 Poisson 累积概率值.查表 A.4 找到所要求的 $F(a)$,只需找到合适的 μ 列和 a 行的交叉处.

例 11.25 某物理学家将一只 Geiger 计数器放在一种放射物附近,记录激发粒子的个数.2 小时内每 10 秒记录一次.从获得的数据,物理学家计算出 10 秒钟(单位)内粒子(成功)的平均率 λ = 每 10 秒 5.5 个粒子.假设这是一个 Poisson 试验,计算 10 秒钟内激发粒子超过 3 个的概率.首先,通过累加各概率计算,然后从附录表 A.4 查找.

解 ☞ 首先计算 10 秒钟内 $x=0, 1, 2, 3$ 个粒子的 Poisson 概率分布.由公式(11.26),

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

对 $\mu=\lambda t=5.5\times1=5.5$,

$$f(0) = \frac{5.5^0 e^{-5.5}}{0!} = \frac{e^{-5.5}}{1} = 0.00409$$

$$f(1) = \frac{5.5^1 e^{-5.5}}{1!} = \frac{5.5\times e^{-5.5}}{1} = 0.02248$$

$$f(2) = \frac{5.5^2 e^{-5.5}}{2!} = \frac{30.25\times e^{-5.5}}{2} = 0.06181$$

$$f(3) = \frac{5.5^3 e^{-5.5}}{3!} = \frac{166.375\times e^{-5.5}}{6} = 0.11332$$

通过累加计算概率为

$$P(X>3) = 1 - P(X\leq 3) = 1 - F(3)$$

这里

$$F(3) = \sum_{x\leq 3}\left[f(x) = \frac{\mu^x e^{-\mu}}{x!}\right]$$

$$\begin{aligned}F(3) &= f(0)+f(1)+f(2)+f(3)\\ &= 0.00409+0.02248+0.06181+0.11332\\ &= 0.20170\end{aligned}$$

从而

$$P(X>3) = 1-0.20170 = 0.79830,\text{ 或 } 0.798$$

现在,从附录表 A.4 查找这个概率值.需要在列($\mu=5.5$)和行($a=3$)的交叉处找到 $F(3)$.交叉处的值是 0.202.得到的解和上面一样,

$$P(X > 3) = 1 - 0.202 = 0.798$$

11.31　Poisson 分布作为二项分布的近似

Poisson 概率分布可用于近似二项概率分布，条件是：在二项试验中随机出现的成功是**稀有事件**，其中 n“大”而 p“小”．我们不讨论这两个分布之间的关系，只简单的给出以下事实：

二项分布 $f(x)=\binom{n}{x}p^xq^{n-x}$ 有两个参数 n 和 p，并且均值 $\mu=np$．如果 np 在 n 增加而 p 减少的过程中保持不变，则当 n 趋近无穷而 p 趋近 0 时，二项分布趋近均值为 $\mu=np$ 的 Poisson 分布 $f(x)=\dfrac{\mu^x e^{-\mu}}{x!}$．

对众多的统计学家，不存在完全一致的“大”和“小”．但是，一个通常的规则是：如果 $n\geqslant 20$ 且 $p\leqslant 0.05$，Poisson 近似是好的；如果 $n\geqslant 100$ 且 $p\leqslant 0.01$，则 Poisson 近似的效果极好．

例 11.26　一家大型工厂聘用了 100 名新员工并进行一项培训．根据以前的上千名培训者情况，项目经理估计将有 4% 的培训者不能完成培训．分别用二项和 Poisson 分布计算恰有 6 人不能完成培训的概率．

解 ☞　如果认为在各次独立试验中 p 保持不变，则这是一个二项试验，其中：成功 = 不能完成的培训者，$n=100$，$p=0.04$，$q=1-0.04=0.96$．于是，由二项概率函数［公式(11.2)］

$$f(x)=\binom{n}{x}p^xq^{n-x}$$

$$f(6)=\binom{100}{6}(0.04)^6(0.96)^{94}$$

$$=(1192052400)(0.0000000041)(0.021552)$$

$$=0.1053$$

由于 $n\geqslant 100$ 且 $p\leqslant 0.05$，可以使用二项分布的 Poisson 近似．于是，对 $\mu=n\times p=100\times 0.04=4$ 由公式(11.26)

$$f(x)=\frac{\mu^x e^{-\mu}}{x!}$$

$$f(6)\approx\frac{4^6e^{-4}}{6!}=\frac{(4096)(0.018316)}{720}=0.1042$$

习 题 解 答

二项概率分布

11.1　下面哪些是二项试验？

(a) 大量研究表明，药物 A 对某种皮肤病的治愈率为 64%．假设你是一名皮肤病专家，将药物 A 临床用于 10 个这样的病人，并记录治愈的病人数．

(b) 总体为 150000 个家庭，从中随机抽出 150 个作为样本并电话询问这些家庭的收入．

解 ☞　(a) 如果以下事实成立：(1) 药物 A 对 10 个病人的治愈率 64% 保持不变，(2) 各次药物治疗相互独立，(3) 患该皮肤病的总体 N 必须很大．则可以接受这项研究是一个二项试验，其中：成功 = 治愈，$p=0.64$，$q=1-0.64=0.36$，二项试验是治疗 10 个病人，二项变量是治愈的病人数．

(b) 这不是一个二项试验，因为不能将结果分为成功和失败两类．

11.2　下面哪些是二项试验？

(a) 某工厂有一台生产电钻部件的机器，取出连续生产的 30 个部件进行检验，记录不合格品的个数．

(b) 在一次选举中,从注册的50000个选举人中随机抽出100人进行调查,记录愿支持候选人A而不支持候选人B的人数.

解 ☞ (a) 这不是一个二项试验,因为各次试验相互不独立.如果机器开始生产不合格品,则随后生产的每一部件很可能也是不合格品.

(b) 由于 $n \leqslant 0.05N[n=(100/50000)N=0.002N]$,故可以认为在各次独立试验中 p 保持不变.所以,这是一个二项试验,其中:成功 = 支持A, p = 50000人中支持A的未知百分比,$q=1-p$,多重Bernoulli试验是调查100个选举人,二项变量是支持A的人数.

11.3 下面哪些是二项试验?

(a) 根据过去经验,轮胎检验员期望批量为1000个某种轮胎的批的不合格品率为2%.从某一批中,检验员取出20个作为随机样本,并记录样本中不合格品个数.

(b) 从与(a)相同的一批中,检验员不是如(a)取出20个轮胎作为样本并记录不合格品个数,而是随机地逐个地取出轮胎,并在第一次发现不合格品时记录已取出的轮胎个数.

解 ☞ (a) 由于 $n \leqslant 0.05N$,可以认为在各次独立试验中 p 保持不变,因此这是一个二项试验.

(b) 不是一个二项试验,因为试验次数 n 不固定.

11.4 习题11.3中的轮胎检验员收到1000个轮胎的一个批,不合格品率为2%.他取出一个容量为5的随机样本,并记录样本中的不合格轮胎个数.由公式(11.2),确定二项变量:不合格品个数的概率分布.并计算变量的均值,方差和标准差.

解 ☞ 所求分布在表11.7中给出,并且在下面给出均值,方差和标准差的计算式.

表 11.7

不合格品数 x	概率 $f(x)=\binom{n}{x}p^x q^{n-x}$
0	$f(0)=\binom{5}{0}(0.02)^0(0.98)^{5-0}=0.9039208$
1	$f(1)=\binom{5}{1}(0.02)^1(0.98)^{5-1}=0.0922368$
2	$f(2)=\binom{5}{2}(0.02)^2(0.98)^{5-2}=0.0037648$
3	$f(3)=\binom{5}{3}(0.02)^3(0.98)^{5-3}=0.0000768$
4	$f(4)=\binom{5}{4}(0.02)^4(0.98)^{5-4}=0.0000008$
5	$f(5)=\binom{5}{5}(0.02)^5(0.98)^{5-5}=0.0000000$
$\sum$	1.0000000

$$\mu=np=(5)(0.02)=0.10$$

$$\sigma^2=npq=(5)(0.02)(0.98)=0.098$$

$$\sigma=\sqrt{\sigma^2}=\sqrt{0.098}=0.313$$

11.5 对上册例10.5的二项试验:掷3次硬币,随机变量是得到正面的次数,3次投掷中得到2次正面的概率是

$$f(2)=P(H_1 \cap H_2 \cap T_3)+P(H_1 \cap T_2 \cap H_3)+P(T_1 \cap H_2 \cap H_3)=3(0.125)=0.375$$

用二项概率函数[公式(11.2)]重新计算,然后解释两个结果相同的原因.

解 ☞ 对于该试验：成功＝得到正面，失败＝得到反面，$p=\frac{1}{2}$，$q=\frac{1}{2}$，$n=3$. 由公式(11.2)，

$$f(x)=\binom{n}{x}p^x q^{n-x}$$

$$f(2)=\binom{3}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^{3-2}$$

$$=\frac{3!}{2!(3-2)!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^1=3\left(\frac{1}{2}\right)^3=3(0.125)=0.375$$

为理解二项概率函数给出与原来计算结果相同的原因，将原来计算变换为二项形式，即

$$f(2)=P(H_1\cap H_2\cap T_3)+P(H_1\cap T_2\cap H_3)+P(T_1\cap H_2\cap H_3)$$

$$=(p\times p\times q)+(p\times q\times p)+(q\times p\times p)$$

可以化简为

$$f(2)=(p^2q)+(p^2q)+(p^2q)=3(p^2q)$$

显然，是二项概率函数的两部分：二项系数是 3，p^xq^{n-x} 项是 p^2q.

11.6 用公式(11.4)，(11.6)和(11.7)计算表 11.2 中二项分布的均值，方差和标准差.

解 ☞ $\mu=np=(7)\left(\frac{1}{2}\right)=3.5$

$\sigma^2=npq=(7)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)=1.75$

$\sigma=\sqrt{npq}=\sqrt{1.75}=1.3229$，或 1.32

11.7 用公式(11.4)，(11.6)和(11.7)计算表 11.1 中二项分布的均值，方差和标准差.

解 ☞ $\mu=np=(5)\left(\frac{1}{6}\right)=0.8333$ 或 0.83

$\sigma^2=npq=(5)\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)=0.69444$ 或 0.694

$\sigma=\sqrt{npq}=\sqrt{0.69444}=0.83333$ 或 0.833

11.8 用 Pascal 三角形计算展开式 $(q+p)^7$ 和 $(q+p)^8$ 的二项系数.

解 ☞ 二项系数如下得到：首先将 $n=6$ 行(例 11.7)加到图 11-1 中的三角形，然后用 11.8 节的计算方法. $n=6$，$n=7$ 和 $n=8$ 的系数在图 11-4 给出.

n									
6	1	6	15	20	15	6	1		
7	1	7	21	35	35	21	7	1	
8	1	8	28	56	70	56	28	8	1

图 11-4

11.9 对二项变量：掷 7 次硬币出现正面的次数，用表 11.2 中二项分布计算出现以下结果的概率：(a)至多 5 次正面，(b)多于 2 次但至多 5 次正面，(c)2，3，4，5 次正面.

解 ☞ (a) $F(5)=P(X\leqslant 5)=f(0)+f(1)+f(2)+f(3)+f(4)+f(5)$

$=0.0078+0.0547+0.1641+0.2734+0.2734+0.1641=0.9375$

(b) 对离散随机变量，上册公式(10.4)表明

$$P(a<X\leqslant b)=F(b)-F(a)$$

从习题 11.8 知 $F(2)=0.2266$；而且由(a)，$F(5)=0.9375$. 因此，

$$P(2<X\leqslant 5)=F(5)-F(2)=0.9375-0.2266=0.7109$$

(c) 对离散随机变量，上册公式(10.5)表明

$$P(a \leqslant X \leqslant b) = F(b) - F(a) + f(a)$$

因此,

$$P(2 \leqslant X \leqslant 5) = F(5) - F(2) + f(2) = 0.9375 - 0.2266 + 0.1641 = 0.8750$$

11.10 对二项变量:掷7次硬币出现正面的次数,用习题11.9的结果和表11.2中二项分布计算出现以下结果的概率:

(a) 多于2次但少于5次正面,(b)至少2次但少于5次正面.

解 ☞ (a)对离散随机变量,上册公式(10.6)表明

$$P(a < X < b) = F(b) - F(a) - f(b)$$

因此,

$$P(2 < X < 5) = F(5) - F(2) - f(5) = 0.9375 - 0.2266 - 0.1641 = 0.5468$$

(b) 对离散随机变量,上册公式(10.7)表明

$$P(a \leqslant X < b) = F(b) - F(a) + f(a) - f(b)$$

因此,

$$P(2 \leqslant X < 5) = F(5) - F(2) + f(2) - f(5)$$

$$= 0.9375 - 0.2266 + 0.1641 - 0.1641 = 0.7109$$

11.11 对二项变量:掷7次硬币出现正面的次数,通过附录表A.3,使用二项累积概率计算出现以下结果的概率:

(a) 至多2次正面,(b)至多3次正面,(c)3次正面.

解 ☞ (a)查表A.3找到行($n=7$, $a=2$)和列($p=0.50$)的交叉处,

$$F(2) = P(X \leqslant 2) = 0.2266$$

这与例11.8的结果一致.

(b) 查表A.3找到行($n=7$, $a=3$)和列($p=0.50$)的交叉处,

$$F(3) = P(X \leqslant 3) = 0.5000$$

(c) $f(3) = P(2 < X \leqslant 3)$.

用上册公式(10.4),

$$P(2 < X \leqslant 3) = F(3) - F(2)$$

因此,由(a)和(b)的结果,

$$f(3) = 0.5000 - 0.2266 = 0.2734$$

结果与表11.2给出的$f(3)$值相同(近似到4位小数).

11.12 对二项变量:掷7次硬币出现正面的次数,使用附录表A.3和习题11.11的结果计算出现以下结果的概率:(a)至少3次正面,(b)多于3次正面,(c)多于2次但少于5次正面.

解 ☞ (a) $P(X \geqslant 3) = 1 - P(X \leqslant 2) = 1 - F(2)$

由习题11.11(a)知 $F(2) = 0.2266$,因此

$$P(X \geqslant 3) = 1 - 0.2266 = 0.7734$$

(b) $P(X > 3) = 1 - P(X \leqslant 3) = 1 - F(3)$.

由习题11.11(a)知 $F(3) = 0.5000$,因此

$$P(X > 3) = 1 - 0.5000 = 0.5000$$

(c) 上册公式(10.6)表明

$$P(a < X < b) = F(b) - F(a) - f(b)$$

因此,

$$P(2 < X < 5) = F(5) - F(2) - f(5)$$

由习题 11.11(a)知 $F(2) = 0.2266$. 查表 A.3, 行($n=7$, $a=4$)和列($p=0.50$)的交叉处是 $F(4)=0.7734$. 行($n=7$, $a=5$)和列($p=0.50$)的交叉处是 $F(5)=0.9375$.

因此,

$$f(5) = F(5) - F(4) = 0.9375 - 0.7734 = 0.1641$$

从而

$$P(2 < X < 5) = 0.9375 - 0.2266 - 0.1641 = 0.5468$$

这与习题 11.10(a)的结果一致.

11.13 应聘某大型公司实验室工作的求职者要求完成由 8 个多项选择题构成的一份基础化学方面的试题. 每个问题有 4 个选项. 如果求职者没有化学知识而只是随机回答, 则: (a)他们平均能答对多少个问题? (b)某个随机答题的求职者至多答对比(a)中平均个数多 1 个的概率是多少?

解☞ (a)如果求职者随机答题, 则这是一个二项试验. 其中: 成功 = 回答正确, $n=8$ 和 $p=0.25$. 因此, 由公式(11.4), 平均答对个数是

$$\mu = np = (8)(0.25) = 2.0$$

(b) 查附录表 A.3, 对于 $n=8$, $a=3$ 和 $p=0.25$

$$P(X \leqslant 3) = F(3) = 0.8862$$

11.14 用附录表 A.3, 对习题 11.13 的应聘考试, 计算: (a)如果求职者至少答对 6 个选择题方能考虑雇用, 则一个随机答题的求职者被考虑的概率是多少? (b)一个随机答题的求职者答对 2,3 或 4 个问题的概率是多少?

解☞ (a) $P(X \geqslant 6) = 1 - F(5)$

查表 A.3, 对于 $n=8$, $a=5$ 和 $p=0.25$

$$F(5) = 0.9958$$

因此,

$$P(X \geqslant 6) = 1 - 0.9958 = 0.0042$$

(b) 由上册公式(10.5)

$$P(2 \leqslant X \leqslant 4) = F(4) - F(2) + f(2)$$

查表 A.3, 对于($n=8$, $a=1$, $p=0.25$), ($n=8$, $a=2$, $p=0.25$)和($n=8$, $a=4$, $p=0.25$),

$$F(1) = 0.3671,\ F(2) = 0.6785,\ F(4) = 0.9727$$

于是,

$$f(2) = F(2) - F(1) = 0.6785 - 0.3671 = 0.3114$$

因此,

$$P(2 \leqslant X \leqslant 4) = 0.9727 - 0.6785 + 0.3114 = 0.6056$$

多项概率分布

11.15 某城市的注册选民有 40% 是民主党人, 35% 是共和党人, 其余 25% 是非两党人士. 如果随机对 10 个选民进行调查, 10 个选民都是共和党人的概率是多少?

解☞ 10 个选民只是城市注册选民的一小部分, 因此调查是一个多项试验, 其中: $n=10$, $A_1=$ {民主党人}, $A_2=$ {共和党人}, $A_3=$ {非两党人士}, $X_1=$ 民主党人数 $=x_1$, $X_2=$ 共和党人数 $=x_2$, $X_3=$ 非两党人数 $=x_3$, $p_1=0.40$, $p_2=0.35$, $p_3=0.25$. 由公式(11.13),

$$f(x_1, x_2, x_3) = \frac{n!}{x_1!x_2!x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

$$f(0, 10, 0) = \frac{10!}{0!10!0!}(0.40)^0(0.35)^{10}(0.25)^0$$

$$= (1)(1)(0.0000276)(1) = 0.000028$$

11.16 对习题 11.15 中 10 个注册选民的随机调查,由公式(11.13)计算概率:(a)2 个民主党人,2 个共和党人,6 个非两党人士,(b)5 个民主党人和 5 个共和党人.

解 ☞ (a)
$$f(2,2,6) = \frac{10!}{2!\,2!\,6!}(0.40)^2(0.35)^2(0.25)^6$$
$$= \left(\frac{3628800}{2880}\right)(0.16)(0.1225)(0.0002441) = 0.006028$$

(b)
$$f(5,5,0) = \frac{10!}{5!\,5!\,0!}(0.40)^5(0.35)^5(0.25)^0$$
$$= \left(\frac{3628800}{14400}\right)(0.01024)(0.005252)(1) = 0.013553$$

11.17 在交叉培育两种植物之后,某遗传学家根据 Mendel 理论预测,后代总体中(高紫花植物 TP)对(矮紫花植物 SP)对(高白花植物 TW)对(矮白花植物 SW)的比例为 9:3:3:1. 假设他是正确的,如果从后代总体中随机选出 16 个种子进行种植,则每种植物各有 4 株的概率是多少?

解 ☞ 如果可以假设 16 个种子只是现在和将来后代总体的一小部分($n \leqslant 0.05N$),则这是一个多项试验,其中:$n=16$, $A_1=\{TP\}$, $A_2=\{SP\}$, $A_3=\{TW\}$, $A_4=\{SW\}$, $X_1=$ TP 个数 $=x_1$, $X_2=$ SP 个数 $=x_2$, $X_3=$ TW 个数 $=x_3$, $X_4=$ SW 个数 $=x_4$, $p_1=9/16$, $p_2=3/16$, $p_3=3/16$, $p_4=1/16$. 由公式(11.13),

$$f(x_1,x_2,x_3,x_4) = \frac{n!}{x_1!\,x_2!\,x_3!\,x_4!}p_1^{x_1}p_2^{x_2}p_3^{x_3}p_4^{x_4}$$

$$f(4,4,4,4) = \frac{16!}{4!4!4!4!}\left(\frac{9}{16}\right)^4\left(\frac{3}{16}\right)^4\left(\frac{3}{16}\right)^4\left(\frac{1}{16}\right)^4$$
$$= (63063000)(0.100113)(0.001236)(0.001236)(0.000015)$$
$$= 0.000145$$

11.18 对习题 11.17 中种植的 16 个种子:(a)有 9TP,3SP,3TW 和 1SW 的概率是多少? (b)这个分布的期望和参数各是多少?

解 ☞ (a)由公式(11.13),

$$f(9,3,3,1) = \frac{16!}{9!3!3!1!}\left(\frac{9}{16}\right)^9\left(\frac{3}{16}\right)^3\left(\frac{3}{16}\right)^3\left(\frac{1}{16}\right)^1$$
$$= (1601600)(0.005638)(0.006592)(0.006592)(0.0625)$$
$$= 0.02452$$

(b) 用公式(11.14),期望是

$$E(X_1) = \mu_1 = np_1 = 16\left(\frac{9}{16}\right) = 9 \quad E(X_2) = \mu_2 = np_2 = 16\left(\frac{3}{16}\right) = 3$$

$$E(X_3) = \mu_3 = np_3 = 16\left(\frac{3}{16}\right) = 3 \quad E(X_4) = \mu_4 = np_4 = 16\left(\frac{1}{16}\right) = 1$$

参数是(见 11.17 节)

$$n = 16,\ p_1 = \frac{9}{16},\ p_2 = \frac{3}{16},\ p_3 = \frac{3}{16},\ p_4 = \frac{1}{16}$$

超几何概率分布

11.19 对例 11.17 的选举委员的试验,计算经理选出 3 个男性员工的概率.

解 ☞ 对这个试验:X 是超几何随机变量,记录 3 次选出的 M 的个数($x=0,1,2,3$),$N_T=$ 员工人数 $=10$,$n=$ 样本容量 $=3$,$N_S=$ 员工中 M 的个数 $=6$,$N_F=$ 员工中 W 的个数 $=4$,$x=$ 样本中 M 的个数,$n-x=$ 样本中 W 的个数. 将这些值代入公式(11.17),

$$f(x)=P(X=x)=\frac{\left[\dfrac{N_S!}{x!(N_S-x)!}\right]\left[\dfrac{N_F!}{(n-x)![N_F-(n-x)]!}\right]}{\left[\dfrac{N_T!}{n!(N_T-n)!}\right]}$$

$$f(3)=\frac{\left[\dfrac{6!}{3!(6-3)!}\right]\left[\dfrac{4!}{(3-3)![4-0]!}\right]}{\left[\dfrac{10!}{3!(10-3)!}\right]}=\frac{20}{120}=0.1667$$

11.20　一只碗中放有 20 粒除颜色外别无差异的石子：10 粒为红色，10 粒为绿色. 闭上眼睛随机取出 10 粒石子，取出后不放回. 计算在取出的 10 粒石子中有 4 粒红色的概率.

解☞　由于是无放回抽样，试验不是多重 Bernoulli 试验，因此本问题必须用超几何分布解决. 对这个试验：X 是超几何随机变量，记录 10 次选取中 R 的个数（$x=0,1,2,\cdots,10$），$N_T=20$，$n=10$，$N_S=10$，$N_F=10$，x＝样本中 R 的个数，$n-x$＝样本中 G 的个数. 因此，由公式(11.16)，

$$f(x)=\frac{\dbinom{N_S}{x}\dbinom{N_F}{n-x}}{\dbinom{N_T}{n}}$$

$$f(4)=\frac{\dbinom{10}{4}\dbinom{10}{6}}{\dbinom{20}{10}}=\frac{(210)(210)}{184756}=0.2387$$

11.21　一家工厂收到批量为 40 的多批保险丝. 接收部门从每一批中随机检验 4 根保险丝，只要任何一根为不合格品则拒绝该批. 如果某一批确有 10% 的不合格品，则抽检的 4 根保险丝中有 1 根为不合格品的概率是多少？

解☞　这是一个超几何试验，其中：X＝不合格品数，$N_T=40$，$n=4$，$N_S=4$，$N_F=36$，$x=1$，$n-x=3$. 因此，由公式(11.16)，

$$f(x)=\frac{\dbinom{N_S}{x}\dbinom{N_F}{n-x}}{\dbinom{N_T}{n}}$$

$$f(1)=\frac{\dbinom{4}{1}\dbinom{36}{3}}{\dbinom{40}{4}}=\frac{(4)(7140)}{91390}=0.3125$$

11.22　对习题 11.21 中 4 根保险丝的检验，该批被接受的概率是多少？

解☞　在 11.11 节，二项分布被用于解决批验收抽样问题. 但是，如果二项分布不合适［无放回抽样且($n/N=0.10$)>0.05］，诸如这里，则可用超几何分布计算在给定接收数 a 下批的接受概率 P_a. 因此，对 $a=0$，由公式(11.16)，

$$P_a=P(X=0)=\frac{\dbinom{4}{0}\dbinom{36}{4}}{\dbinom{40}{4}}=\frac{(1)(58905)}{91390}=0.6445$$

于是，即使不合格品率为 10% 的一批仍有 64% 的接受概率.

11.23　对习题 11.21 和 11.22 的工厂，现在是以批量为 50 收到保险丝. 工厂决定从每一批中随机检验 6 根保险丝，如果多于 1 根是不合格品则拒绝该批. 如果批的实际不合格品

率是 12%，则 P_a 是多少？

解 ☞　这是一个超几何试验，其中：X = 不合格品数，$N_T = 50$，$n = 6$，$N_S = 6$，$N_F = 44$，$x = 0$ 或 1，$n - x = 6$ 或 5. 将公式(11.16)代入上册公式(10.2)，得到

$$P_a = P(X \leqslant a) = F(a) = \sum_{x \leqslant a} \left[f(x) = \frac{\binom{N_S}{x}\binom{N_F}{n-x}}{\binom{N_T}{n}} \right]$$

对本问题

$$P_a = P(X \leqslant 1) = F(1) = \frac{\binom{6}{0}\binom{44}{6}}{\binom{50}{6}} + \frac{\binom{6}{1}\binom{44}{5}}{\binom{50}{6}}$$

$$= \frac{(1)(7059052)}{15890700} + \frac{(1)(1086008)}{15890700}$$

$$= 0.442 + 0.4101 = 0.8543$$

于是，即使不合格品率为 12% 的一批仍有 85% 的接受概率.

11.24　习题 11.23 的检验保险丝试验的超几何分布的均值，方差和参数是多少？

解 ☞　由公式(11.18)和(11.19)，

$$\mu = \frac{nN_S}{N_T} = \frac{6 \times 6}{50} = 0.72$$

$$\sigma^2 = \frac{nN_SN_F(N_T - n)}{N_T^2(N_T - 1)} = \frac{6 \times 6 \times 44(50 - 6)}{50^2(50 - 1)} = 0.569$$

参数是(见 11.22 节)$N_T = 50$，$n = 6$，$N_S = 6$.

11.25　一组油漆工漆了 38 座房子. 油漆工作完成情况是：14 个房主非常满意(记为 V)，16 个房主一般满意(记为 M)，8 个房主不满意(记为 D). 油漆公司老板决定随机调查其中的 12 个房主. 调查包括 4V，2M 和 6D 的概率是多少？

解 ☞　这是一个超几何试验，可以使用超几何概率函数的推广形式[公式(11.21)] 试验中：X_1 = 样本中 V 的个数 = $x_1 = 4$，X_2 = 样本中 M 的个数 = $x_2 = 2$，X_3 = 样本中 D 的个数 = $x_3 = 6$，$N_T = 38$，N_1 = 总体中 V 的个数 = 14，N_2 = 总体中 M 的个数 = 16，N_3 = 总体中 D 的个数 = 8，n = 12. 因此，

$$f(x_1, x_2, x_3) = \frac{\binom{N_1}{x_1}\binom{N_2}{x_2}\binom{N_3}{x_3}}{\binom{N_T}{n}}$$

$$f(4, 2, 6) = \frac{\binom{14}{4}\binom{16}{2}\binom{8}{6}}{\binom{38}{12}} = \frac{(1001)(120)(28)}{(2707475148)} = 0.00124$$

Poisson 分布

11.26　对例 11.25 中 Geiger 计数器的 Poisson 试验，记录粒子的平均比例是 λ = 每 10 秒 5.5 个粒子. 使用附录表 A.4 和上册 10.7 节的原理，计算 10 秒内记录有 6 个粒子的概率.

解 ☞　由上册公式(10.4)，

$$P(5 < X \leqslant 6) = F(6) - F(5) = f(6)$$

对 $\mu=5.5$ 查表 A.4,得 $F(5)=0.529$ 和 $F(6)=0.686$.因此,

$$f(6)=0.686-0.529=0.157$$

11.27 对例 11.26 的关于 Geiger 计数器的 Poisson 试验,用习题 11.26 的结果和上册 10.7 节的原理,计算 10 秒内有以下记录结果的概率:(a) 至少 6 个粒子,(b)6,7 或 8 个粒子.

解 ☞　(a) $P(X\geqslant 6)=1-P(X\leqslant 5)=1-F(5)$

从习题 11.26 知,$F(5)=0.529$.因此,

$$P(X\geqslant 6)=1-0.529=0.471$$

(b) 由上册公式(10.5),

$$P(6\leqslant X\leqslant 8)=F(8)-F(6)+f(6)$$

从习题 11.26 知,$F(6)=0.686$ 和 $f(6)=0.157$;对 $\mu=5.5$ 查表 A.4, 得到 $F(8)=0.894$,因此,

$$P(6\leqslant X\leqslant 8)=0.894-0.686+0.157=0.365$$

11.28 Poisson 分布经常用于商业中的**库存控制系统**.诸如,一家海鲜餐厅的经理想用 Poisson 分布确定每天需要的活龙虾的数量,已知每天顾客平均定购 7 只龙虾.假设这是一个 Poisson试验,则:(a)如果某天餐厅有 9 只龙虾,则龙虾定购量多于 9 只的概率是多少?(b)如果某天经理希望有大于 95%的把握满足需求,则他需要多少只龙虾?

解 ☞　(a) 问题是:给定 $\mu=\lambda t=7\times 1=7$, 计算 $P(X>9)$. 已知

$$P(X>9)=1-P(X\leqslant 9)=1-F(9)$$

对 $\mu=7$ 查表 A.4,得 $F(9)=0.830$.因此,手头没有足够龙虾的概率是

$$P(X>9)=1-0.830=0.170$$

(b) 问题是:已知 $\mu=7$,下式中的 a 是多少?

$$P(X\leqslant a)=F(a)>0.950$$

对 $\mu=7$ 查表 A.4,得 $F(11)=0.947$ 和 $F(12)=0.973$,如果经理想有大于 95% 的把握,则应该有 12 只龙虾.

11.29 一家大型工厂聘用了 100 人并进行一项培训.根据以前的上千个培训者情况,项目经理估计将有 4%的培训者不能完成培训课程.计算 $n=100$ 个培训者中至少有 5 个不能完成培训的概率.

解 ☞　这个问题不能使用表 A.3,因为表 A.3 不包括较大的 n 值,大部分二项表也如此.但是,可以用表 A.4 求解,此时 n 只出现在$\mu=n\times p=100\times 0.04=4$ 中.须解出

$$P(X\geqslant 5)=1-P(X\leqslant 4)=1-F(4)$$

对 $\mu=4$ 查表 A.4,得 $F(4)=0.629$.于是,

$$P(X\geqslant 5)\approx 1-0.629=0.371$$

11.30 经过多次试验,发现一种疫苗对某种疾病的免疫有效率为 98.4%.用附录表 A.4 和上册 10.7 节的原理,计算:在现在注射疫苗的 125 人中,有以下无效结果的概率:(a)3 人,(b)多于 3 人.

解 ☞　(a)如果认为这是一个二项试验,则可以用 Poisson 近似方法[$(n=125)>100$; $(p=0.016)<0.05$]求解.于是,由上册公式(10.4),

$$P(2<X\leqslant 3)=F(3)-F(2)=f(3)$$

对 $\mu=n\times p=125\times 0.016=2$ 查表 A.4,得 $F(3)=0.857$ 和 $F(2)=0.677$.因此,

$$f(3)\approx 0.857-0.677=0.180$$

(b) $P(X>3)=1-P(X\leqslant 3)=1-F(3)\approx 1-0.857=0.143$

11.31 习题 11.28 中的餐厅平均每天有 280 个顾客.如果经理决定针对顾客流量而不是时间

段进行库存控制,即针对当前的200个顾客的需求而不是每天的需求,则这200个顾客定购龙虾的数量为 $x=8, 9, 10, 11$ 只的概率是多少?

解 ☞　如果认为这是一个二项试验,则可以用Poisson近似方法[($n=200$)>100;($p=7/280=0.025$)<0.05]求解.因此,将 $\mu=n\times p=200\times0.025=5$ 代入公式(11.26),

$$f(x)=\frac{\mu^x e^{-\mu}}{x!}$$

$$f(8)\approx\frac{5^8 e^{-5}}{8!}=\frac{(390625)(0.006738)}{40320}=0.06528$$

$$f(9)\approx\frac{5^9 e^{-5}}{9!}=\frac{(1953125)(0.006738)}{362880}=0.03627$$

$$f(10)\approx\frac{5^{10} e^{-5}}{10!}=\frac{(9765625)(0.006738)}{3628800}=0.01813$$

$$f(11)\approx\frac{5^{11} e^{-5}}{11!}=\frac{(48828125)(0.006738)}{39916800}=0.00824$$

补充习题

二项概率分布

11.32　下面哪些是二项试验?

(a) 从一副52张扑克牌中随机取出10张,每次选取后进行以下操作:观测扑克是否为红色,放回扑克,重新洗牌.记录10次选取中红色的张数.

(b) 同(a)中试验相同,只是每次取出后不放回.

答案:(a)是二项试验,(b)不是二项试验

11.33　下面是一个二项试验:掷一粒骰子5次,记录出现4点的次数.对于这个二项试验,说明:成功,失败,p,q 和二项随机变量分别是什么?

答案:成功=掷一次骰子观测到4点,失败=观测到其它结果(1,2,3,5,6点),$p=1/6$,$q=1-1/6=5/6$,二项随机变量是掷5次骰子出现4点的次数.

11.34　一个职业棒球队员投球已达上千次,成功的概率是80%.在连续9次投球中,计算出现以下成功结果的概率:(a)至少5次,(b)至多5次,(c)少于5次.

答案:(a) 0.9804,(b) 0.0856,(c) 0.0196

11.35　50个杂志推销员每天晚上从花名册上随机选出10个人并打电话.每次电话的任务是完成一个杂志订阅.从上千次这种电话,杂志商了解到只有15%成功,即得到一个新的订阅.对于每10次电话:(a)公司期望得到的订阅数的均值和标准差是多少?(b)订阅数不多于2份的概率是多少?

答案:(a) $\mu=1.50$,$\sigma=1.129$;(b) 0.8202

11.36　对于习题11.35的50个推销员,每10次电话得到多于1份新订阅的概率是多少?

答案:0.4557

多项概率分布

11.37　多项试验是:从一副标准的52张扑克牌中选取6张,每次选取后放回并重新洗牌.如果每次试验的互斥且完备的可能事件是 $A_1=\{A, 2, 3, 4, 5\}$,$A_2=\{6, 7, 8\}$,$A_3=\{9, 10\}$,$A_4=\{J, Q, K\}$,则(a) $f(0, 0, 0, 6)$,(b) $f(2, 1, 3, 0)$ 是多少?

答案:(a) 0.000151,(b) 0.007458

11.38　对习题11.37的选取扑克牌试验的概率分布,试问:(a) 期望,(b)参数是多少?

答案:(a) $E(X_1)=2.31$,$E(X_2)=1.38$,$E(X_3)=0.92$,$E(X_4)=1.38$;(b) $n=6$,$p_1=\frac{5}{13}$,$p_2=\frac{3}{13}$,$p_3=\frac{2}{13}$,$p_4=\frac{3}{13}$

11.39　某学院的院长要派8名学生参加一个全国性会议.学院有350名一年级学生,250名二年级学生,200名三年级学生和200名四年级学生.学生的名字完全混合排列,院长从中随机选出8个.由于8个只是学生总数的一小部分($n\leqslant0.05N$),故这是一个多项试验.试问:从每个年级选出2名学生的概率是

多少?

答案:0.0309

11.40　一个野生生物学家在某小岛上设置捕鼠器连续捕获并随即放走 15 只老鼠(即,有放回抽样),岛上生活的不同鼠类的个数是:A 类 25 只,B 类 20 只,C 类 15 只.假设老鼠随机进入捕鼠器,则容量为 15 的一个样本包括每类老鼠各 5 只的概率是多少?

答案:0.0382

超几何概率分布

11.41　某学院的院长要派 8 名学生参加一个全国性会议.申请者为 12 名二年级学生和 8 名四年级学生.他们的名字完全混合放在一顶帽子里,院长从中随机取出 8 个.试问:选出 4 个二年级学生的概率是多少?

答案:0.275

11.42　一个野生生物学家设置捕鼠器从 60 只老鼠(36 只雄性和 24 只雌性)的一个总体中捕获了 10 只.捕获的老鼠不放回总体,而是被带回试验室以便研究其行为.假设老鼠随机进入捕鼠器,则容量为 10 的一个样本中有 5 只雄性的概率是多少?

答案:0.213

11.43　有 16 个男生和 14 个女生选修代数课.假设所有学生均来上课且随机进入教室,试问:最初进入教室的 5 个学生都是女生的概率是多少?

答案:0.0140

11.44　一亩森林有 100 棵成年松树,对真菌感染它们不表现出任何特征.然而,以往的研究表明,30% 的松树有早期感染,而且这只能通过检查内部组织检测到.假设数字 30% 是正确的,如果伐木工人砍倒 10 棵松树并检查内部组织,发现有 5 棵树受感染的概率是多少?

答案:0.0996

Poisson 概率分布

11.45　对习题 11.26 的关于 Geiger 计数器的 Poisson 试验,用附录表 A.4 和上册 10.7 节的原理,计算 10 秒内出现以下记录结果的概率:(a)多于 2 个但少于 6 个粒子,(b)多于 6 个但最多 9 个粒子.

答案:(a)0.441, (b)0.260

11.46　对于习题 11.30 的疫苗试验,多次试验表明,疫苗对某种疾病的免疫有效率为 98.4%.用附录表 A.4 和上册 10.7 节的原理,计算:在随后注射疫苗的 125 个病人中,疫苗出现以下无效结果的概率:(a)多于 0 个但少于 3 个,(b)0,1,2 或 3 个.

答案:如果认为这是一个二项试验,则可以使用 Poisson 近似方法.(a)≈0.542, (b)≈0.857

11.47　在一项艰难的攀岩运动中,夏季每周有 10 个攀岩者不能成功登顶.如果不能成功登顶的人数服从 Poisson分布,则在七月的某两周内有 12 个攀岩者不能成功登顶的概率是多少?

答案:0.0176

11.48　研究某种鸟巢的鸟类学家在该鸟的栖息地沿一条篱笆观察并记录下看到的鸟巢个数.在 1 公里(1000 米)内他记录下 30 个鸟巢.试问:在某 200 米内他发现 2 个鸟巢的概率是多少?

答案:0.0446

11.49　每小时有 30 辆汽车穿过一条人行道.在某 5 分钟内,没有汽车穿过该人行道的概率是多少?

答案:0.0821

11.50　对习题 11.49 描述的汽车,在某 5 分钟内,穿过人行道的汽车不超过 4 辆的概率是多少?

答案:0.891

第十二章　正态分布和其它连续型概率分布

12.1　连续型概率分布

在第 11 章中，我们讨论了四个离散型概率分布：二项分布，多项分布，超几何分布和 Poisson 分布的具体特征和应用．本章继续讨论三种连续型概率分布：**正态分布，均匀分布和指数分布**．不过，在开始讨论之前，我们先通过例 12.1 回顾一下上册第 10 章已经讲过的连续型分布的一般性质．

例 12.1　对于图 12-1 给出的连续型概率分布，试问：(a) $f(a)$，(b) $P(X=a)$，(c) $F(a)$　(d) $P(a\leqslant X\leqslant b)$，(e) $P(a<X<b)$，(f) $P(-\infty<X<\infty)$，(g) $P(b<X<\infty)$，(h) μ，(i) σ^2 和 σ 各是多少?

解　(a) 图 12-1 显示的连续曲线是对一个概率密度函数（也称为连续型概率分布）的描述．该函数用 $f(x)$ 表示，基于一个可取无限且不可数多个确定值 x（见上册，10.4 节）的连续随机变量 X．如果将 $x=a$ 代入 $f(x)$，结果是 $f(a)$．

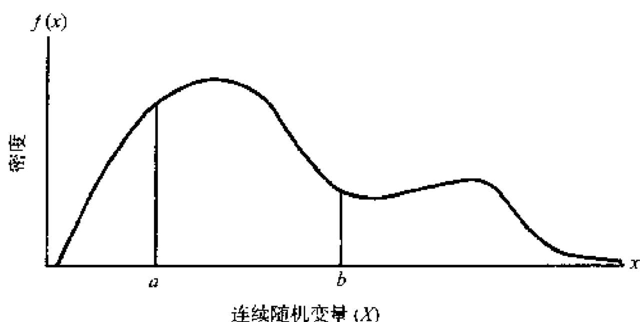


图 12-1

(b) 由上册图 10-3(b) 的性质(4)知，$P(X=x)=0$．这表明 X 取确定值 a 的概率是零，即 $P(X=a)=0$．

(c) 由上册 10.8 节知，对 $X=a$，连续随机变量的累积分布函数是[上册，公式(10.8)]

$$F(a)=P(X\leqslant a)=\int_{-\infty}^{a}f(x)dx$$

表明，X 取值小于或等于 a 的概率在图 12-1 中是 a 点垂线左边的曲线下方的面积．

(d) 由上册图 10-3(b) 的性质(6)知，

$$P(a\leqslant X\leqslant b)=\int_{a}^{b}f(x)dx$$

表明，X 在区间 a 到 b 上取值的概率在图 12-1 中是以曲线，X 轴以及 a 和 b 点上的垂线为边界的面积．由上册公式(10.9)也知，

$$P(a\leqslant X\leqslant b)=F(b)-F(a)$$

(e) 由上册公式(10.9)知，

$$P(a<X\leqslant b)=P(a\leqslant X\leqslant b)=P(a<X<b)$$
$$-P(a\leqslant X<b)=F(b)-F(a)$$

(f) 由上册图 10-3(b) 的性质(7)，对于连续型概率分布

$$P(-\infty<X<\infty)=\int_{-\infty}^{\infty}f(x)dx=P(S)=1.00$$

表明，由于概率分布包括连续随机变量 X 的所有可能结果的概率，故曲线下方的总面积为 1.00．

(g) 由(c)和(f)，

$$P(b < X < \infty) = 1 - F(b)$$

(h) 从上册 10.10 节知,连续随机变量 X[或连续型概率分布 $f(x)$]的均值或期望是公式(10.14):

$$E(X) = \mu_x = \mu = \int_{-\infty}^{\infty} xf(x)dx$$

(i) 从上册 10.13 节知,连续随机变量 X[或连续型概率分布 $f(x)$]的方差是公式(10.24):

$$\mathrm{Var}(X) = \sigma_x^2 = \sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$$

标准差是

$$\sigma = \sqrt{\sigma^2}$$

12.2　正态概率分布和正态概率密度函数

正态概率分布(或**正态分布**)是统计学中最重要的理论上的连续型分布,因为:(1)在推断性统计学的发展中它扮演着中心角色,(2)现实世界中许多随机变量表现出来的频数(或相对频数)分布非常接近正态分布,(3)它能方便的用于近似许多其它的概率分布,诸如二项分布和 Poisson 分布(见 12.12 和 12.13 节). 正态分布在上册一些章节中已经出现过,并且还将在本卷的剩余章节中经常出现.

同离散型概率分布(见第十一章)一样,所有的连续型概率分布也都用唯一确定的概率函数来定义. 用来定义正态概率分布的函数称为**正态概率密度函数**或**正态概率函数**(见上册,10.4 节),对于可取确定值 $X = x$ 的连续随机变量 X,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2},\ -\infty < x < \infty \tag{12.1}$$

这里 $e = 2.71828\cdots$[见上册,习题 1.23(b)]

$\pi = 3.14159\cdots$(见上册,1.12 节)

$\mu = E(X)$ = 正态分布的均值

$\sigma^2 = E[(X-\mu)^2]$ = 正态分布的方差

$\sigma = \sqrt{\sigma^2}$ = 正态分布的标准差

$-\infty < x < \infty$ 表明该函数对一切实数有定义

图 12-2 给出正态分布的典型图形. 横轴表示连续随机变量 X 的确定值 x,竖轴表示正态概率密度函数 $f(x)$ 的确定值. 图中的光滑曲线称为**正态曲线**,是对足够多的 $X = x$ 值(横坐标,见上册,1.20 节)计算 $f(x)$ 值(纵坐标,见上册,1.20 节)而得到. 可以看到,曲线酷似钟

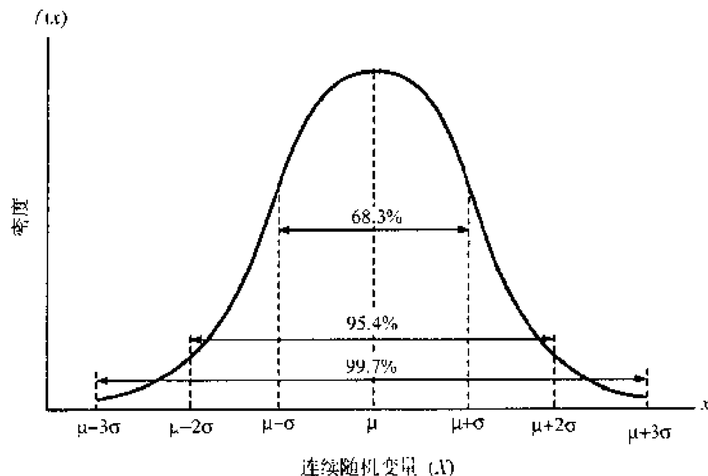


图 12-2

形,并且关于均值 μ 上的垂线完全对称. 于是,曲线下方的面积有 50%处于这条垂线的左边,50%处于右边. 曲线向着负无穷和正无穷连续延伸,即在两个方向上越来越接近横轴但永不相交,不过这在图形中无法显示. 对于任何连续型概率分布,曲线下方的总面积是 1.0,而且图形也给出落在区间 $\mu \pm \sigma$(68.3%), $\mu \pm 2\sigma$(95.4%)和 $\mu \pm 3\sigma$(99.7%)上的面积的百分数. 在上册 7.16 节,这些百分数作为**经验法则**曾经给出过,而且,它们对习题 12.5 的正态分布也成立,这在稍后说明.

正态分布的公式最早在 1733 年由法国数学家 Abraham de Moivre(1667—1754)发表,当时他用这个公式近似二项分布. 尽管法国数学家和天文学家 Pierre Simon de Lsplace(1749—1827)延续了 de Moivre 的工作,但是公认的最早对正态分布的性质和用途进行实质性探索的是德国数学家,天文学家和物理学家 Karl Friedrich Gauss(1777—1855). 正因为这一点,正态分布也称为 **Gauss 分布**.

如果一个数据集的频数(或相对频数)分布能合理的用正态曲线拟合,则该数据集称为是正态分布的. 正如前面所述(见上册,7.16 节),即使一个经验分布只具有单峰,大概的小丘形并且基本对称,即只是近似于正态分布,也常称为正态分布. 现实世界中许多随机变量有这样的分布:某一男性群体的身高,求职者的行为测试得分,一批甜瓜的重量,一群妇女的心脏血压,等等. 虽然正态曲线对现实世界中的变量非常普遍,但它并不是因为其它曲线的"反常态"才称为"常态". 实际上,术语"正态"在这个分布的历史发展早期就已使用,因为它与另一条所谓的**正态误差曲线**相类似.

12.3　正态概率分布族

同第十一章的离散型概率分布一样,连续型概率分布也是一些分布族,讨论中涉及的某一具体分布由它的一个或多个参数决定. 从公式(12.1)可以看到,正态分布有两个参数:均值 μ 和方差 σ^2(一些统计著作也说成是 μ 和 σ).

例 12.2　对参数为 $\mu=0$ 和 $\sigma^2=1$ 的正态分布用公式(12.1)计算 $X=1$ 时的 $f(x)$.

解 ☞　将 $\mu=0$ 和 $\sigma^2=1$ 代入公式(12.1),

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$

$$f(1) = \frac{1}{\sqrt{1}\times\sqrt{2\pi}}e^{-(1-0)^2/2\times 1}$$

$$= \frac{1}{\sqrt{2\times 3.14159}}e^{-1/2} = \frac{1}{2.50663}e^{-1/2}$$

$$= 0.39894e^{-1/2}$$

然后,用例 11.22 计算 e^{-x} 的方法,

$$f(1) = (0.39894)(0.60653) = 0.24197$$

12.4　正态分布:均值(μ),中位数($\tilde{\mu}$)和众数的关系

上册 10.6 节指出,连续型概率分布函数是连续随机变量的总体相对频数分布的数学模型. 因此,概率分布和总体分布必须用可以相互比较的统计测量进行刻画:二者都有均值,中位数和众数. 在正态分布情形,两个分布既是单峰的也是对称的;并且,由上册习题 6.29(a)我们知道,对于二者,均值(μ)=中位数($\tilde{\mu}$)=众数.

例 12.3　对于图 12-3 给出的两个正态分布(A 和 B),哪一个有较大的:(a)μ, (b)σ,(c)σ^2?

解 ☞　(a) 在任何正态分布中,均值 μ 指出了分布的中位数在横轴上(连续随机变量)的位置. 因此,在 μ 点对 $f(x)$ 所做的垂线将分布分成镜像对称的两部分(见图 12-2). 据此,从图 12-3 可以看到,B 的均值(μ_B)在 A 的均值(μ_A)的右边. 因为对所有直角 Cartes 坐标系,横轴上原点右端的值是正的且递增(见上

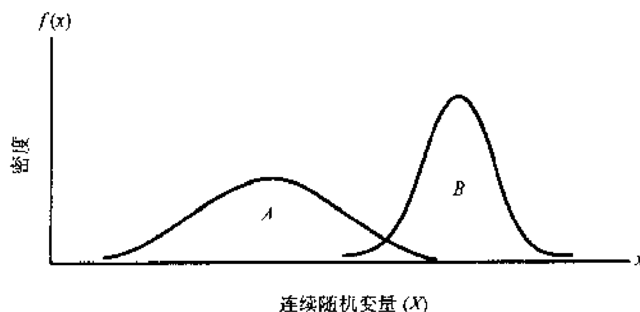


图 12-3

册、1.20 节），于是，可知 $\mu_A < \mu_B$.

(b) 从上册 7.9 节知，标准差是变量取值对均值的离散（散布）程度的度量．根据经验法则（见 12.2 节），正态分布 68.3% 的面积总是落在与均值相距一个标准差的范围之内（$\mu \pm \sigma$）．因此，标准差越小，聚集在均值附近的值越多，使分布有更陡的峰，或有更大的**峰度**．由此，从图 12-3 可知 $\sigma_B < \sigma_A$．

(c) 由上，$\sigma_B^2 \leqslant \sigma_A^2$．

12.5　峰度

不是所有的单峰，对称分布都是正态分布．这些分布与正态的差异在于峰度，即峰的陡峭程度[见例 12.3(b)]．图 12-4 给出三个有不同峰度的分布．中间的分布是一个正态分布，称为**常峰态的**．左边的分布与正态相比有较少和比较扁平的峰，并且尾部较短，称为**低峰态的**．右边的分布，与正态分布相比有较多的峰且在中部取值集中，并且尾部较长，称为**尖峰态的**．

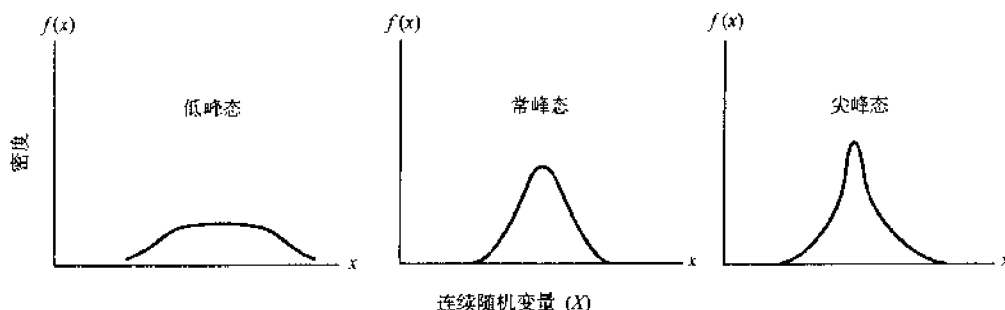


图 12-4

12.6　标准正态分布

标准正态分布（或标准正态概率分布或标准化正态分布或单位正态分布）是有指定参数：$\mu = 0$ 和 $\sigma^2 = 1$ 的正态分布．将这两个参数代入公式(12.1)得到**标准正态概率密度函数或标准正态密度函数**：

$$f(x) = \frac{1}{1\sqrt{2\pi}} e^{-(x-0)^2/2(1)^2} = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{12.2}$$

图 12-5 给出标准正态分布的图形．同任何其它的正态分布一样，这条光滑曲线称为**标准正态曲线**，是通过对足够多的 $X = x$ 值计算密度值[$f(x)$]画出的．注意：由于 $\mu = 0$ 和 $\sigma^2 = \sigma = 1$，该分布总是关于纵（密度）轴对称，并且横轴的刻度总是只显示 -3 到 3 的范围．这是因为，99.7% 的 $f(x)$ 值落在 $\mu - 3\sigma$ 和 $\mu + 3\sigma$ 之间（见 12.2 节），对于标准正态分布这个范围就是 0 ± 3．

对于例 12.2 中讨论的 $\mu = 0$ 和 $\sigma^2 = \sigma = 1$ 的正态分布，现在我们知道，其实就是标准正态分布．当时，计算出 $f(1) = 0.24197$，这个值用划圈的实点标注在图 12-5 的标准正态曲线上．

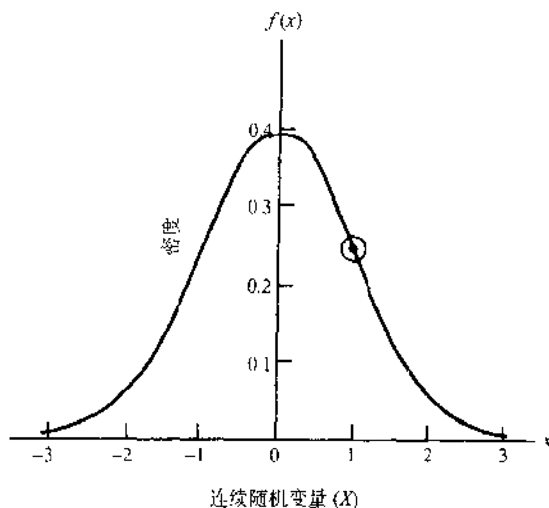


图 12-5

12.7 标准正态分布和标准正态变量之间的关系

上册 7.19 节介绍了**标准得分**(**或正态离差或** z **得分**)[公式(7.46)]:

$$z_i = \frac{x_i - \mu}{\sigma}$$

并且指出,对任何变量 X,样本或总体中的每一个测量 x 都能变换为一个 z 值,这个过程就是常说的对变量进行**标准化**或**正规化**. 得到的 Z 变量称为**标准化变量**. 接着,上册习题 7.37 指出,对于来自标准化变量的总体或任意样本,必然成立:均值为 0($\bar{z}=0$ 或 $\mu_z=0$),标准差(从而方差)为 1($s_z^2=s_z=1$ 或 $\sigma_z^2=\sigma_z=1$).

上述概念可应用于连续随机变量及其概率分布. 如果连续随机变量 X 有一个均值 μ 和方差 σ^2 的正态分布,并且所有 $X=x$ 值均变换为标准得分(或这里是**标准正态离差**),则变换生成一个新的连续随机变量 Z,称为**标准正态变量**,即

$$Z = \frac{X - \mu}{\sigma} \tag{12.3}$$

能取范围 $-\infty < z < \infty$ 内的任何确定值

$$z_i = \frac{x_i - \mu}{\sigma} \tag{12.4}$$

与上册习题 7.37 给出的理由相同,我们指出,标准正态变量的均值总为 0($\mu=0$)且方差和标准差总为 1($\sigma^2=\sigma=1$). 于是,标准正态变量的概率分布就是标准正态分布,形式如公式(12.2)定义:

$$f(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2},\ -\infty < z < \infty \tag{12.5}$$

这表明,通过将连续随机变量 X 进行 Z 变换,任何正态分布都能变换为标准正态分布. 对于有无穷多个的理论正态分布族或者能用正态曲线拟合的经验分布中的任何一个分布,这个结论总成立. 在经验分布情形,Z 变换将原始度量单位(例如,克)转换为**标准单位**. 标准单位是分布的标准差的倍数,标准差用来反映某个指定的 x 值偏离分布均值有多远.

只有当初始的 X 服从正态分布时,X 的 Z 变换才能生成标准正态分布. Z 变换不能把非正态分布(例如,有偏分布)变换为标准正态分布.

12.8 标准正态分布的面积表

由连续型概率分布(见例12.1)的一般性质知:(1) 分布的总面积是1.0,(2) 连续随机变量 X 在区间 a 到 b 上取值的概率是以曲线,X 轴以及 a 和 b 上的垂线为边界的分布中的面积,(3) 对于所考虑的具体的概率函数,这些面积能通过积分方法得到.

尽管对正态分布族的无穷多个分布的任一个都可以使用积分解决面积/概率问题,但无此必要. 因为,任意正态分布都能变换为标准正态分布,所以任意正态分布的面积/概率问题都能通过对公式(12.5),即该分布的密度函数,进行积分运算得到.

对0.00到3.99变化的正的 z 值,附录表A.5(**标准正态分布的面积**)提供了公式(12.5)的积分值:

$$P(0 \leqslant Z \leqslant z) = \int_0^z \left[f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right] dz \qquad (12.6)$$

这个积分结果是以标准正态曲线,Z 轴,$z=0$ 点的垂线和感兴趣的 z 值点的垂线为边界的面积.

例 12.4 用表A.5回答以下问题:(a) 标准正态分布落在 $0 \leqslant z \leqslant 0.45$ 上的面积是多少? (b) 标准正态分布落在 $0 \leqslant z \leqslant 0.45$ 上的比例是多少? (c) 标准正态变量 Z 在区间 $0 \leqslant z \leqslant 0.45$ 上取值的概率是多少?

解 ☞ (a) 曲线下方 $0 \leqslant z \leqslant 0.45$ 上的面积是图12-6显示的标准正态分布的阴影面积. 它以曲线,Z 轴,$z=0$ 点的垂线和 $z=0.45$ 点的垂线为边界. 这个面积能从表A.5找到,首先在左列找到 z 的一个小数(0.4)位置,然后在列标题找到 z 的另一个小数(0.05)位置. 表12.1给出表A.5的一部分,可以看到,在行 $z=0.4$ 和列 $z=0.05$ 的交叉处,面积(划有圆圈)是0.1736.

(b) 因为标准正态分布的总面积是1.0,故表A.5给出的面积就是1.0的一个比例. 于是,标准正态分布落在 $0 \leqslant z \leqslant 0.45$ 上的比例也是0.1736.

(c) 由于(面积=连续型概率分布中的概率),

$$P(0 \leqslant Z \leqslant 0.45) = \int_0^{0.45} \left[f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \right] dz = 0.1736$$

表 12.1

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879

例 12.5 用表A.5计算标准正态分布落在 $-1.69 \leqslant z \leqslant 0$ 上的面积.

解 ☞ 在 $z=0$ 处,密度轴[$f(z)$](见图12-6)将标准正态分布分成对称的两部分. 这表明,如果 $+a$ 和 $-a$ 是分别落在 $z=0$ 的右边和左边的 Z 轴上的点,并且与 $z=0$ 等距,则曲线在 $0 \leqslant z \leqslant a$ 和 $-a \leqslant z \leqslant 0$ 上的面积相等且彼此对称. 因此,表A.5给出的 $0 \leqslant z \leqslant a$ 上的面积也是 $-a \leqslant z \leqslant 0$ 上的面积. 这样,对于本问题,表A.5给出曲线在 $0 \leqslant z \leqslant 1.69$ 上的面积值是0.4545,这也是曲线在 $-1.69 \leqslant z \leqslant 0$ 上的面积. -1.69 和0之间的面积是图12-7显示的标准正态分布的阴影面积.

例 12.6 根据例12.4和例12.5得到的解答,试问:$P(-1.69 \leqslant Z \leqslant 0.45)$ 是多少?

解 ☞ $P(-1.69 \leqslant Z \leqslant 0.45)$ 在图12-8显示的标准正态分布中是曲线在 $-1.69 \leqslant z \leqslant 0.45$ 上的阴影面积. 它是例12.4和例12.5得到的面积的和:

$$P(-1.69 \leqslant X \leqslant 0.45) = 0.4545 + 0.1736 = 0.6281$$

12.9 利用 Z 变换计算任意正态分布的概率

在数学上可以证明

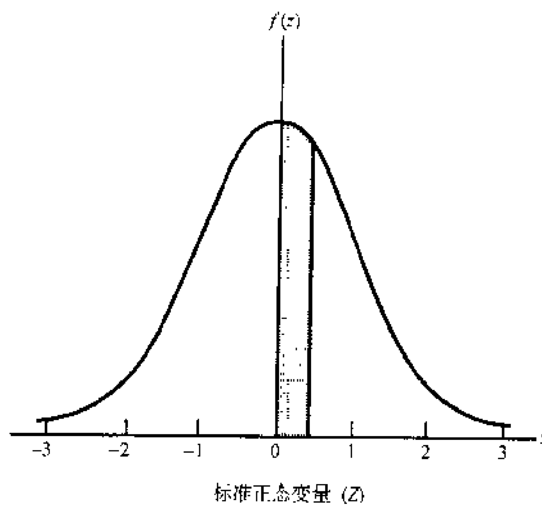


图 12-6

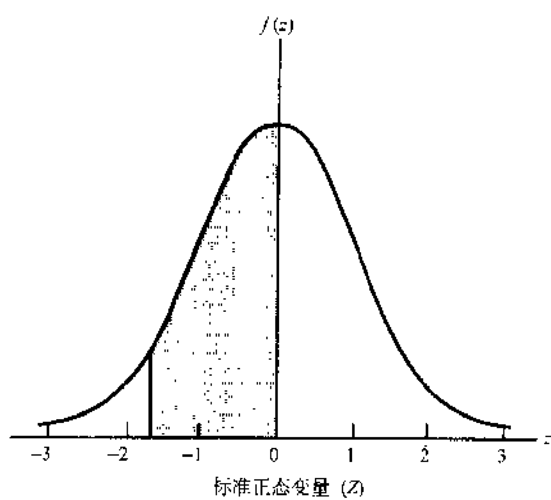


图 12-7

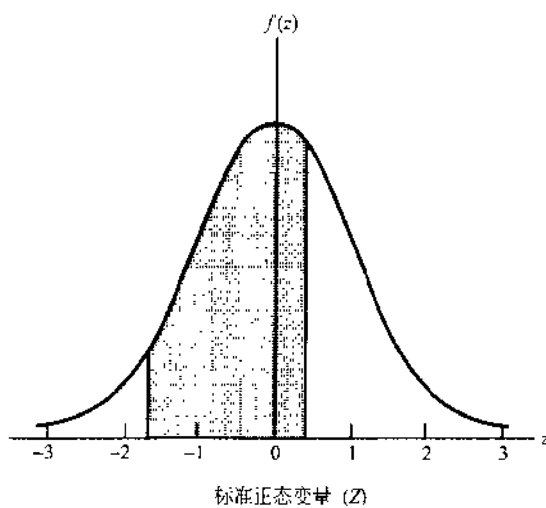


图 12-8

$$P(a \leqslant X \leqslant b) = P\left(\frac{a-\mu}{\sigma} \leqslant Z \leqslant \frac{b-\mu}{\sigma}\right) = P(z_a \leqslant Z \leqslant z_b) \quad (12.7)$$

表明，对任何均值为 μ，方差为 σ^2 的正态分布，它的随机变量 X 在区间 $a \leqslant x \leqslant b$ 上取值的概率与标准正态变量 Z 在区间

$$\left(\frac{a-\mu}{\sigma} \leqslant Z \leqslant \frac{b-\mu}{\sigma}\right) = (z_a \leqslant Z \leqslant z_b)$$

上取值的概率相同.

例 12.7　某生理学家想了解冬眠对地鼠体重的影响. 他设计了一个试验，在夏末和早春称量 1000 只成年地鼠的体重. 在夏末的称量中，他发现地鼠体重是正态分布的，均值是 400g，标准差是 100g. 试问：在夏末某只地鼠的体重在 350g 和 450g 之间的概率是多少？

解 ☞　解答这个问题，需要将 $P(350\text{g} \leqslant X \leqslant 450\text{g})$ 变换为 $P(z_a \leqslant Z \leqslant z_b)$ 的形式. 首先，用公式(12.4)将 350g 和 450g 变换为 z 得分：

$$z_a = \frac{350\text{g} - 400\text{g}}{100\text{g}} = -0.5 \quad z_b = \frac{450\text{g} - 400\text{g}}{100\text{g}} = 0.5$$

接着，将 z 得分代入公式(12.7)

$$P(a \leqslant X \leqslant b) = P(z_a \leqslant Z \leqslant z_b)$$

$$P(350\text{g} \leqslant X \leqslant 450\text{g}) = P(-0.5 \leqslant Z \leqslant 0.5)$$

查附录表 A.5.找到 $P(0 \leqslant Z \leqslant 0.5) = 0.1915$. 因为正态曲线的对称性(见例 12.5)，$P(-0.5 \leqslant Z \leqslant 0) = P(0 \leqslant Z \leqslant 0.5)$，因此

$$P(-0.5 \leqslant Z \leqslant 0.5) = 2(0.1915) = 0.3830$$

一只地鼠体重在 350g 和 450g 之间的概率是 0.3830.

注意：在例 12.7，1000 只地鼠的样本被视为一个总体，在公式中用到了总体参数 μ 和 σ 这样做是因为，本章使用的标准正态方法要求统计信息或者是直接来自总体，或者是来自足够大的样本(如例 12.7)，由此给出总体分布的形状，均值和标准差的准确信息.

12.10　单尾概率

许多统计方法是处理单尾概率，即感兴趣的只是概率分布的**上尾(右尾)**或**下尾(左尾)**的面积. 如果用 α(小写希腊字母 alpha)表示上尾面积，对随机变量 X，x_α 是满足

$$P(X > x_\alpha) = \alpha \quad (12.8)$$

的一个定值，换句话说，这表明在曲线下方 x_α 右边的面积(概率)是 α. 如果 X 是一个正态分布的连续变量，则在数学上可以证明

$$P(X > x_\alpha) = P\left(Z > \frac{x_\alpha - \mu}{\sigma}\right) = P(Z > z_\alpha) = \alpha \quad (12.9)$$

这个概率是图 12-9 所示的标准概率分布中的阴影面积.

类似地，如果用 α 表示下尾面积，对变量 X，$x_{1-\alpha}$ 是满足

$$P(X < x_{1-\alpha}) = \alpha \quad (12.10)$$

的一个定值. 换句话说，在曲线下方 $x_{1-\alpha}$ 左边的面积(概率)是 α. (下标 $1-\alpha$ 表示 $x_{1-\alpha}$ 右边的面积是 $1-\alpha$.)如果 X 是一个正态分布的连续变量，在数学上可以证明

$$P(X < x_{1-\alpha}) = P\left(Z < \frac{x_{1-\alpha} - \mu}{\sigma}\right) = P(Z < -z_\alpha) = \alpha \quad (12.11)$$

这个概率是图 12-10 所示的标准正态分布中的阴影面积.

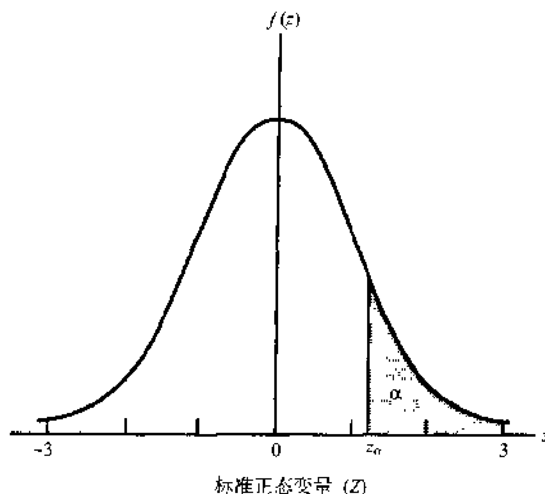


图 12-9

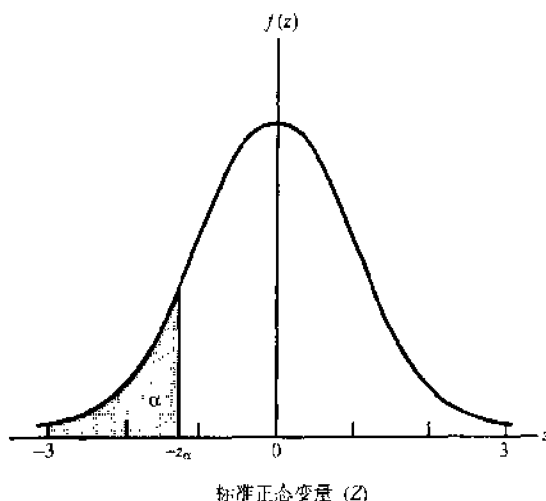


图 12-10

例 12.8 如果正态分布变量 X 有参数 $\mu=14.5$ 和 $\sigma=2.1$,对 $\alpha=0.05$,试问:(a) z_α 和 x_α,(b) $-z_\alpha$ 和 $x_{1-\alpha}$ 各是多少?

解 ☞ (a)由公式(12.9),

$$P(Z>z_\alpha)=\alpha=0.05$$

于是,为求 z_α,需要查附录表 A.5 找到 z 值,使得对应的面积最接近 $0.5-0.05=0.45$

$$P(0\leqslant Z\leqslant z_\alpha)=0.45$$

我们发现,表中"最接近"0.45 的面积不止一个,而是有两个面积都与 0.45 只相差 0.0005:对 0.4495,有 $z=1.64$;对 0.4505,有 $z=1.65$。因此,使用两个 z 值的平均作为 z_α

$$z_\alpha=z_{0.05}=\frac{1.64+1.65}{2}=1.645$$

现在,为求 $x_\alpha=x_{0.05}$,整理公式(12.4),得到

$$x_i=\mu+z_i\sigma \tag{12.12}$$

用 z_α 和 x_α 替换 z_i 和 x_i,

$$x_\alpha=\mu+z_\alpha\sigma$$

$$\begin{aligned}x_{0.05}&=\mu+z_{0.05}\sigma=14.5+(1.645\times2.1)\\&=14.5+3.4545=17.9545,\text{ 或 }18.0\end{aligned}$$

(b) 由公式(12.11)知

$$P(Z < -z_\alpha) = \alpha = 0.05$$

因为标准正态分布的对称性，由(a)知

$$-z_\alpha = -z_{0.05} = -1.645$$

为求 $x_{1-\alpha}$，用 $-z_\alpha$ 和 $x_{1-\alpha}$ 替换公式(12.12)中的 z_i 和 x_i：

$$x_{1-\alpha} = \mu + (-z_\alpha\sigma)$$

$$x_{1-0.05} = \mu + (-z_{0.05}\sigma) = 14.5 + (-1.645 \times 2.1)$$

$$= 14.5 - 3.4545 = 11.0455 \text{ 或 } 11.0$$

12.11　双尾概率

对于**双尾概率**，感兴趣的概率 α 在概率分布的两个尾部被平分，而不是像单尾概率那样所有的 α 落在一个尾部．如只用 $\alpha/2$ 表示每个尾部的相等面积，对随机变量 X，确定值 $x_{\alpha/2}$ 和 $x_{1-\alpha/2}$ 满足

$$P(X > x_{\alpha/2}) = \alpha/2 \tag{12.13}$$

$$P(X < x_{1-\alpha/2}) = \alpha/2 \tag{12.14}$$

$$P(x_{1-\alpha/2} \leqslant X \leqslant x_{\alpha/2}) = 1 - \alpha \tag{12.15}$$

换句话说，这些公式表明，在曲线下方 $x_{\alpha/2}$ 右边的面积是 $\alpha/2$，在 $x_{1-\alpha/2}$ 左边也是 $\alpha/2$．因此，$x_{\alpha/2}$ 和 $x_{1-\alpha/2}$ 之间的面积是 $1-\alpha/2-\alpha/2=1-\alpha$．如果 X 是正态分布的连续变量，则也有

$$P(X > x_{\alpha/2}) = P\left(Z > \frac{x_{\alpha/2}-\mu}{\sigma}\right) = P(Z > z_{\alpha/2}) = \frac{\alpha}{2} \tag{12.16}$$

$$P(X < x_{1-\alpha/2}) = P\left(Z < \frac{x_{1-\alpha/2}-\mu}{\sigma}\right) = P(Z < -z_{\alpha/2}) = \frac{\alpha}{2} \tag{12.17}$$

$$P(x_{1-\alpha/2} \leqslant X \leqslant x_{\alpha/2}) = P(-z_{\alpha/2} \leqslant X \leqslant z_{\alpha/2}) = 1 - \alpha \tag{12.18}$$

对标准正态分布，这些面积如图 12-11 所示．

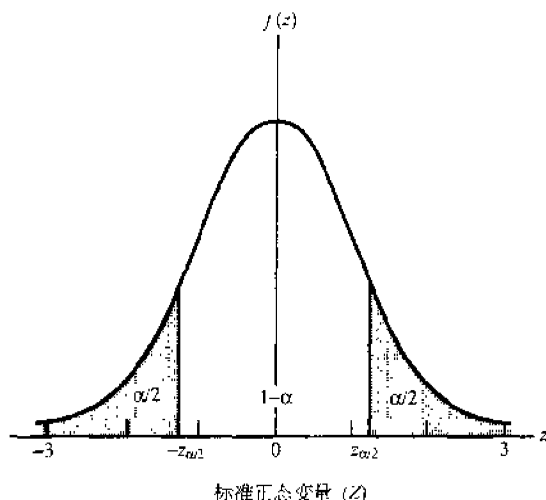


图 12-11

例 12.9　对 $\mu=14.5$ 和 $\sigma=2.1$ 的正态分布变量 X，如果 $\alpha=0.05$，则 $z_{\alpha/2}$，$-z_{\alpha/2}$，$x_{\alpha/2}$ 和 $x_{1-\alpha/2}$ 各是多少？

解 ☞　由公式(12.16)，

$$P(Z > z_{\alpha/2}) = \alpha/2 = 0.05/2 = 0.025$$

因此,为求 $z_{\alpha/2} = z_{0.05/2}$,必须查附录表 A.5 找到 z 值,使得对应的面积最接近 0.5 − 0.025 = 0.475

$$P(0 \leqslant Z \leqslant z_{\alpha/2}) = 0.475$$

这个"最接近"的面积恰是 0.4750,从而 $z_{\alpha/2} = z_{0.05/2} = 1.96$. 由于标准正态分布的对称性,有 $-z_{\alpha/2} = -z_{0.05/2} = -1.96$. 下面求 $x_{\alpha/2}$,将 $x_{\alpha/2}$ 和 $z_{\alpha/2}$ 代入公式(12.12),

$$x_{\alpha/2} = \mu + z_{\alpha/2}\sigma$$

$$x_{0.05/2} = \mu + z_{0.05/2}\sigma = 14.5 + (1.96 \times 2.1)$$

$$= 14.5 + 4.116 = 18.616, \text{或 } 18.6$$

为求 $x_{1-\alpha/2}$,将 $x_{1-\alpha/2}$ 和 $-z_{\alpha/2}$ 代入公式(12.12),

$$x_{1-\alpha/2} = \mu + (-z_{\alpha/2}\sigma)$$

$$x_{1-0.05/2} = \mu + (-z_{0.05/2}\sigma) = 14.5 + (-1.96 \times 2.1)$$

$$= 14.5 - 4.116 = 10.384 \text{ 或 } 10.4$$

12.12　二项分布的正态近似

计算二项分布的某个概率,需要对二项式展开 $(q+p)^n$(见 11.7 节)的相关项相加,这是一个繁琐的过程,而且对较大的 n 值或许得有一台计算机. 然而,11.31 节指出,在一定条件下(即,当 $n \geqslant 20$ 且 $p \leqslant 0.05$ 时),可以使用 Poisson 分布对二项概率进行近似. 类似地,在一定条件下正态分布也是二项分布的一个良好近似,从而可用于计算二项分布的概率. 由于这样得到的概率只是对二项概率真实值的近似,故正态分布的这种应用称为**二项分布的正态近似**.

只有所考虑的二项分布与正态分布相似(即近似),即对称且具有钟形,使用正态近似方法才是合适的,在以下两个条件下,上述结果成立:

(1) p 值(成功的概率)越接近 0.5(见图 11-2),二项分布越对称.

(2) 对任何 p 值,无论距离 0.5 多远,如果 p 值不变而 n(试验次数,或样本容量)增大,得到的 $\mu = np$ 和 $\sigma^2 = npq$ 的二项分布在形状上越来越近似一个 $\mu = np$ 和 $\sigma^2 = npq$ 的正态分布.

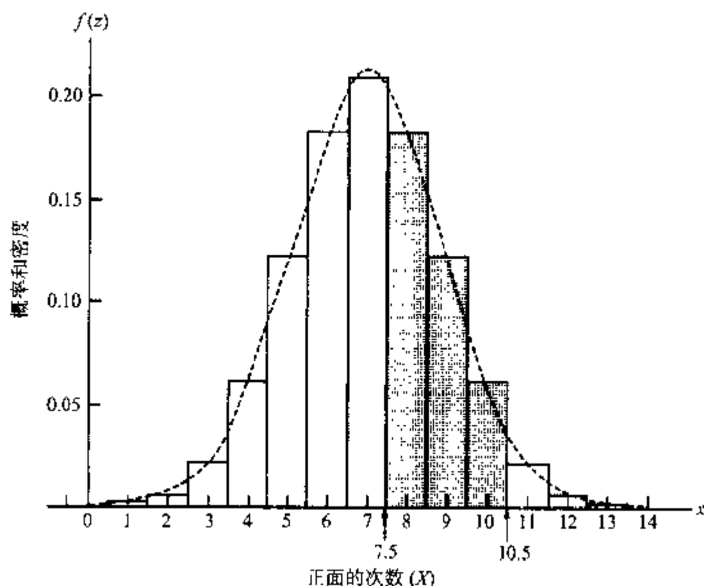


图 12-12

因为有条件(2)，当 n“充分大”时，可以将二项分布看做 $\mu=np$ 和 $\sigma^2=npq$ 的正态分布，然后用标准正态方法近似二项概率的真实值.

对“足够接近”0.5 或 n“充分大”不存在绝对的规则，但是有许多常用的规则说明何时使用正态近似是合适的. 其中，在许多统计学著作可以找到的一个规则是：当 np 和 nq 两者都大于或等于 5($np\geqslant5$ 且 $nq\geqslant5$)时，可以使用二项分布的正态近似.

这个规则更严格的说法是 np 和 nq 都大于 5($np>5$ 且 $nq>5$).

例 12.10 对二项变量：掷一枚硬币 14 次出现正面的次数，计算在 14 次投掷中出现 8,9 或 10 次正面的概率. 使用：(a) 二项分布方法，(b) 正态近似方法.

解 ☞ (a) 用二项分布方法求解这个问题，必须确定变量的二项概率分布. 首先，用二项式定理找到二项展开式$(q+p)^{14}$(见 11.7 节)的各项

$$(q+p)^{14}=q^{14}+14q^{13}p+91q^{12}p^2+364q^{11}p^3+1001q^{10}p^4+2002q^9p^5$$
$$+3003q^8p^6+3432q^7p^7+3003q^6p^8+2002q^5p^9+1001q^4p^{10}$$
$$+364q^3p^{11}+91q^2p^{12}+14qp^{13}+p^{14}$$

然后，将出现一次正面的概率($p=\frac{1}{2}$)和出现一次反面的概率($q=\frac{1}{2}$)代入展开式，给出图 12-12 中的直方图所示的分布. 在这个直方图中，离散二项变量被当做是“连续的”(见上册习题 5.9)，对不同的正面次数，概率值用数字上方的矩形的高或面积(高×1.0)表示. 例如，8 次正面的概率既是数字 8 上方矩形的高(0.1833)又是该矩形的面积(0.1833×1.0=0.1833). 由于这是一个概率分布，矩形的总面积是 1.0，而且，对这个二项概率分布，均值是 $\mu=np=(14)\left(\frac{1}{2}\right)=7$，方差是 $\sigma^2=npq=(14)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)=3.5$，标准差是 $\sigma=\sqrt{npq}=\sqrt{3.5}=1.8708$.

由上册 8.6 节的性质 4 知

$$P(A_1\cup A_2\cup\cdots\cup A_k)=P(A_1)+P(A_2)+\cdots+P(A_k)$$

由于二项变量是离散变量，本问题要求的概率是

$$P(8\leqslant X\leqslant10)=P[(X=8)\cup(X=9)\cup(X=10)]$$
$$=P(X=8)+P(X=9)+P(X=10)$$

由二项式展开$(q+p)^{14}$

$$P(X=8)=3003q^6p^8=3003\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^8=0.1833$$

$$P(X=9)=2002q^5p^9=2002\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^9=0.1222$$

$$P(X=10)=1001q^4p^{10}=1001\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^{10}=0.0611$$

因此，

$$P(8\leqslant X\leqslant10)=0.1833+0.1222+0.0611=0.3666$$

这个概率是图 12-12 的直方图中的阴影面积.

(b) 由于 $np=nq=7$，用正态近似是合适的. 为说明这一点，连同二项分布，图 12-12 还描出了参数为 $\mu=np=7$和 $\sigma^2=npq=3.5$ 的正态分布(虚线).

二项概率 $P(8\leqslant X\leqslant10)=0.3666$ 的正态近似值似乎就是正态曲线在区间 $8\leqslant x\leqslant10$ 上的面积，但是这种说法忽略了事实：二项直方图中这个概率的面积延伸到了区间 $7.5\leqslant x\leqslant10.5$ 上(见图 12-12). 为了用连续型分布近似离散型分布，必须将每个离散值看做一个区间——离散值就是使用的每个测量区间(见上册，2.10 节)的中点，而每个测量区间上下各延伸了 0.5. 这样，在这个问题中，正态曲线在区间 $7.5\leqslant x\leqslant10.5$ 上的面积对测量区间的下界 8 和上界 10 做了延伸. 将离散值加上或减去 0.5 称为**连续性修正**(或**修正为连续**或**连续性的半单位修正**).

为了对 $\mu=np=7$, $\sigma=\sqrt{npq}=\sqrt{3.5}=1.8708$ 的正态分布计算 $P(7.5\leqslant X\leqslant10.5)$，首先用公式(12.7)

将正态概率值变换为标准正态概率值.

$$P(a \leqslant X \leqslant b) = P\left(\frac{a-\mu}{\sigma} \leqslant Z \leqslant \frac{b-\mu}{\sigma}\right) = P(z_a \leqslant Z \leqslant z_b)$$

$$\begin{aligned} P(7.5 \leqslant X \leqslant 10.5) &= P\left(\frac{7.5-np}{\sqrt{npq}} \leqslant Z \leqslant \frac{10.5-np}{\sqrt{npq}}\right) \\ &= P\left(\frac{7.5-7}{1.8708} \leqslant Z \leqslant \frac{10.5-7}{1.8708}\right) \\ &= P(0.27 \leqslant Z \leqslant 1.87) \end{aligned}$$

查附录表 A.5 知

$$P(0 \leqslant Z \leqslant 0.27) = 0.1064 \quad P(0 \leqslant Z \leqslant 1.87) = 0.4693$$

因此

$$P(0.27 \leqslant Z \leqslant 1.87) = 0.4693 - 0.1064 = 0.3629$$

将(a)中二项真实值和正态近似值进行比较,可以看到它们在前二位小数上是相同的:

$$P(8 \leqslant Z \leqslant 10) = 0.3666 \approx P(0.27 \leqslant Z \leqslant 1.87) = 0.3629$$

12.13 Poisson 分布的正态近似

正态分布能用于近似 Poisson 分布. Poisson 分布的参数是 $\mu = \lambda t$(见 11.28 节). 可以证明,随着 $\mu = \lambda t$ 增加,Poisson 分布接近 $\mu = \sigma^2 = \lambda t$ 的正态分布. 因此,只要 λt"足够大",就可以将 Poisson 分布看做是 $\mu = \sigma^2 = \lambda t$ 的正态分布,并且 $\sigma = \sqrt{\lambda t}$. 然后可以用标准正态方法计算面积(概率)值. 因为这样得到的概率值只是 Poisson 概率真实值的近似,所以正态分布的这种应用称为 **Poisson 分布的正态近似**.

对 λt"足够大"不存在绝对的规则,但是同二项分布一样,有许多常用规则说明何时使用正态近似是合适的. 其中,在许多统计学著作可以看到的一个规则是

当 λt 大于或等于 5($\lambda t \geqslant 5$) 时,可以使用 Poisson 分布的正态近似.

这个规则的更严格的形式要求,λt 必须大于或等于 10($\lambda t \geqslant 10$).

例 12.11 在 11.26 节,电缆制造商已经确知,以 4 米为一单位的电缆,每 4 米有 $\lambda = 4.0$ 个缺陷. 对 $t = 1.5$ 单位的电缆,计算有 6,7 或 8 个缺陷的概率,使用:(a) Poisson 分布方法,(b) 正态近似方法

解 ☞ (a) 假设这是一个 Poisson 试验(见 11.26 节),则需要求出 Poisson 分布中的 $P(6 \leqslant X \leqslant 8)$,这里 $\mu = \sigma^2 = \lambda t$ = (每 4 米 4.0 个缺陷) × (1.5 单位) = 每 6 米 6.0 个缺陷. 由上册公式(10.5):

$$P(a \leqslant X \leqslant b) = F(b) - F(a) + f(a)$$

$$P(6 \leqslant X \leqslant 8) = F(8) - F(6) + f(6)$$

和公式(10.4):

$$P(5 < X \leqslant 6) = F(6) - F(5) = f(6)$$

计算这个概率. 对 $\mu = 6.0$,查附录表 A.4 找到:$F(5) = 0.446$, $F(6) = 0.606$ 和 $F(8) = 0.847$. 从而:$f(6) = 0.606 - 0.446 = 0.160$, $P(6 \leqslant X \leqslant 8) = 0.847 - 0.606 + 0.160 = 0.401$.

(b) 因为是用连续型分布近似离散型分布,必须做连续性修正[见例 12.10(b)]. 为求 $P(6 \leqslant X \leqslant 8)$,需要在 $\mu = \sigma^2 = \lambda t = 6.0$ 和 $\sigma = \sqrt{6.0} = 2.4495$ 的正态分布中求 $P(5.5 \leqslant X \leqslant 8.5)$:

$$\begin{aligned} P(5.5 \leqslant X \leqslant 8.5) &= P\left(\frac{5.5-\lambda t}{\sqrt{\lambda t}} \leqslant Z \leqslant \frac{8.5-\lambda t}{\sqrt{\lambda t}}\right) \\ &= P\left(\frac{5.5-6.0}{2.4495} \leqslant Z \leqslant \frac{8.5-6.0}{2.4495}\right) \\ &= P(-0.20 \leqslant Z \leqslant 1.02) \end{aligned}$$

查附录表 A.5 得

$$P(0 \leqslant Z \leqslant 0.20) = P(-0.20 \leqslant Z \leqslant 0) = 0.0793$$

$$P(0 \leqslant Z \leqslant 1.02) = 0.3461$$

因此,

$$P(-0.20 \leqslant Z \leqslant 1.02) = 0.0793 + 0.3461 = 0.4254$$

比较 Poisson 真值和正态近似值,可以看到它们相差 0.024:

$$P(6 \leqslant Z \leqslant 8) = 0.401 \approx P(-0.20 \leqslant Z \leqslant 1.02) = 0.425$$

12.14 离散型均匀概率分布

均匀概率分布的特征是它的随机变量 X 的所有取值有相等的概率,虽然本章是讨论连续型分布,但在这里我们还是要对均匀概率分布的离散形式做一描述,作为引入连续形式的预备.

在**离散型均匀概率分布**或**离散型均匀分布**或**离散型矩形分布**中,随机变量 X 可以取从 1 到 k 的任何整数值,而且所有的 k 个值有相等的概率. 将这个相等概率值分配给 k 个可能结果的概率函数,即定义这个分布的概率函数,称为**离散型均匀概率分布**

$$f(x) = \frac{1}{k},\ x = 0,1,\cdots,k \tag{12.19}$$

由于这是离散型概率分布,必然成立:对任何 x 有 $f(x) \geqslant 0$,且 $\sum_x f(x) = 1.00$[见上册,图 10.3(a)]. 离散型均匀分布的均值是

$$\mu = \frac{k+1}{2} \tag{12.20}$$

方差是

$$\sigma^2 = \frac{k^2-1}{12} \tag{12.21}$$

标准差是

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{k^2-1}{12}} \tag{12.22}$$

离散型均匀概率分布是一个分布族,只有一个参数 k.

例 12.12 对于掷骰子试验,随机变量(最终向上的点数)的概率分布是一个离散型均匀概率分布. 对该分布,计算:(a) 概率函数 $f(x)$, (b) 均值 μ, (c) 方差 σ^2,(d) 标准差 σ,(e) 参数,(f) $P(3 \leqslant X < 5)$.

解 (a) $f(x) = \frac{1}{k} = \frac{1}{6}$, $x = 0,1,\cdots,6$,

(b) $\mu = \frac{k+1}{2} = \frac{6+1}{2} = 3.5$,

(c) $\sigma^2 = \frac{k^2-1}{12} = \frac{6^2-1}{12} = \frac{35}{12} = 2.92$,

(d) $\sigma = \sqrt{\sigma^2} = \sqrt{2.92} = 1.71$,

(e) $k = 6$,

(f) $P(3 \leqslant X < 5) = P(X=3) + P(X=4) = f(3) + f(4) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = 0.3333$.

这个概率是如图 12-13 所示的该分布的概率直方图中的阴影面积.

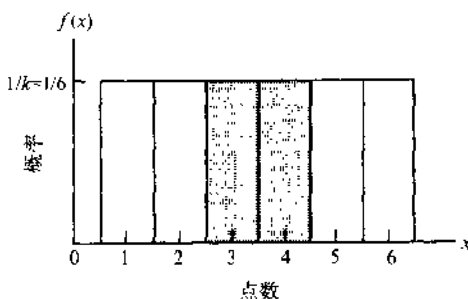


图 12-13

12.15 连续型均匀概率分布

如果连续随机变量 X 可以且只能在区间 $a \leqslant x \leqslant b$ 上任意取值，并且 X 在区间上的概率密度函数 $f(x)$ 是常数(均匀)，而在其它处 $f(x)$ 等于 0，则 X 称为**是均匀分布的**，X 的分布函数称为**连续型均匀概率函数**(或**连续型均匀分布**，或**连续型矩形分布**)．定义这个分布的概率密度函数称为**均匀概率密度函数**，有如下形式：

$$f(x) = \frac{1}{b-a}，当 a \leqslant x \leqslant b 时；f(x) = 0，其它 \tag{12.23}$$

由于这个函数定义了一个连续型概率分布，必然对任何 x 有 $f(x) \geqslant 0$，且 $\int_{-\infty}^{\infty} f(x)dx = 1.00$ [见上册，图 10.3(b)]．连续型均匀分布的均值是

$$\mu = \frac{b+a}{2} \tag{12.24}$$

方差是

$$\sigma^2 = \frac{(b-a)^2}{12} \tag{12.25}$$

标准差是

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{(b-a)^2}{12}} \tag{12.26}$$

连续型均匀分布是一个分布族，有两个参数 a 和 b．

例 12.13 连续随机变量 X 可以在区间 $1 \leqslant x \leqslant 6$ 上任意取值．如果 X 有连续型均匀概率分布，对该分布计算：(a) 概率密度函数 $f(x)$，(b) 均值 μ，(c) 方差 σ^2、(d) 标准差 σ，(e) 参数，(f) $P(3 \leqslant X \leqslant 5)$．

解 (a) $f(x) = \frac{1}{b-a} = \frac{1}{6-1} = \frac{1}{5} = 0.2$，当 $1 \leqslant x \leqslant 6$ 时；$f(x) = 0$，其它，

(b) $\mu = \frac{b+a}{2} = \frac{6+1}{2} = 3.5$，

(c) $\sigma^2 = \frac{(b-a)^2}{12} = \frac{(6-1)^2}{12} = \frac{25}{12} = 2.08$，

(d) $\sigma = \sqrt{\sigma^2} = \sqrt{2.08} = 1.44$，

(e) $a = 1$，$b = 6$，

(f) 只在区间 $a \leqslant x \leqslant b$ 上取值的连续型均匀概率分布的图形是在此区间上高度为 $f(x) = \frac{1}{b-a}$ 的一条水平直线．因为这是一个概率分布，所以分布在 $a \leqslant x \leqslant b$ 上的总面积是 1．本题的分布图形由图 12-14 给出，阴影面积表示 $P(3 \leqslant X \leqslant 5)$．

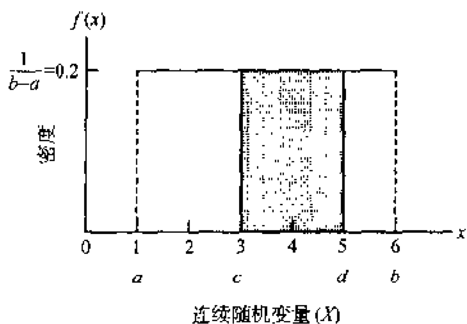


图 12-14

对于连续型分布，概率总是一个区间上的积分；也是指定区间上曲线下方的面积．对连续型均匀分布，就像在图 12-14 看到的一样，面积是以指定区间的长度($5-3=2$)为底，以密度函数[$f(x) = 0.2$]为高的一个矩形．总之，如果区间 $c \leqslant x \leqslant d$ 在指定区间 $a \leqslant x \leqslant b$ 之内，则

$$P(c \leqslant X \leqslant d) = (底)(高) = (d-c)\left(f(x) = \frac{1}{b-a}\right) = \frac{d-c}{b-a}$$

由于 X 是连续变量，区间端点是否包括在区间内无关紧要．因此，

$$P(c \leqslant X \leqslant d) = P(c \leqslant X < d)$$

从而

$$P(c \leqslant X < d) = \frac{d-c}{b-a} = \frac{5-3}{6-1} = \frac{2}{5} = 0.4$$

12.16　指数概率分布

随机变量 X 称为服从**指数分布**，如果对任意 $\lambda > 0$，X 的概率密度函数由**指数概率密度函数**

$$f(x) = \lambda e^{-\lambda x}，当 x \geqslant 0 时；\quad f(x) = 0 \quad 当 x < 0 时 \tag{12.27}$$

给出，这里 $e = 2.71828\cdots$[见上册，习题 1.23(b)]．**指数概率分布**(或**指数分布**)是一个有单参数 λ 的分布族．图 12-15 给出指数分布在 $\lambda = \frac{1}{2}$和 $\lambda = 2$ 时的形式．对指数分布，

$$\mu = \frac{1}{\lambda} \tag{12.28}$$

$$\sigma^2 = \frac{1}{\lambda^2} \tag{12.29}$$

$$\sigma = \frac{1}{\lambda} \tag{12.30}$$

不做推导，我们给出**指数随机变量** X 的**累积分布函数** $F(x) = P(X \leqslant x)$ 的具体形式

$$F(x) = 1 - e^{-\lambda x}，当 x \geqslant 0 时；\quad F(x) = 0 \quad 当 x < 0 时 \tag{12.31}$$

例 12.14　对图 12-15 显示的两个指数分布，确定：均值、方差，标准差，参数和累积分布函数．

解 ☞　对 $\lambda = \frac{1}{2}$ 的分布，

$$\mu = \frac{1}{\lambda} = \frac{1}{1/2} = 2$$

$$\sigma^2 = \frac{1}{\lambda^2} = \frac{1}{(1/2)^2} = 4$$

$$\sigma = \frac{1}{\lambda} = \frac{1}{1/2} = 2$$

参数是 $\lambda = \frac{1}{2}$，累积分布函数是

$$F(x) = 1 - e^{-\lambda x} = 1 - e^{-(1/2)x}$$

对 $\lambda = 2$ 的分布，

$$\mu = \frac{1}{\lambda} = \frac{1}{2}$$

$$\sigma^2 = \frac{1}{\lambda^2} = \frac{1}{2^2} = \frac{1}{4}$$

$$\sigma = \frac{1}{\lambda} = \frac{1}{2}$$

参数是 $\lambda = 2$，累积分布函数是

$$F(x) = 1 - e^{-\lambda x} = 1 - e^{-2x}$$

12.17　指数分布和 Poisson 分布的关系

很可能你已经注意到，正常数 λ(小写希腊字母 lambda)既出现在指数概率密度函数[公式(12.27)]又出现在 Poisson 概率函数[公式(11.22)]．两个分布有相同的参数 λ 是因为有如下

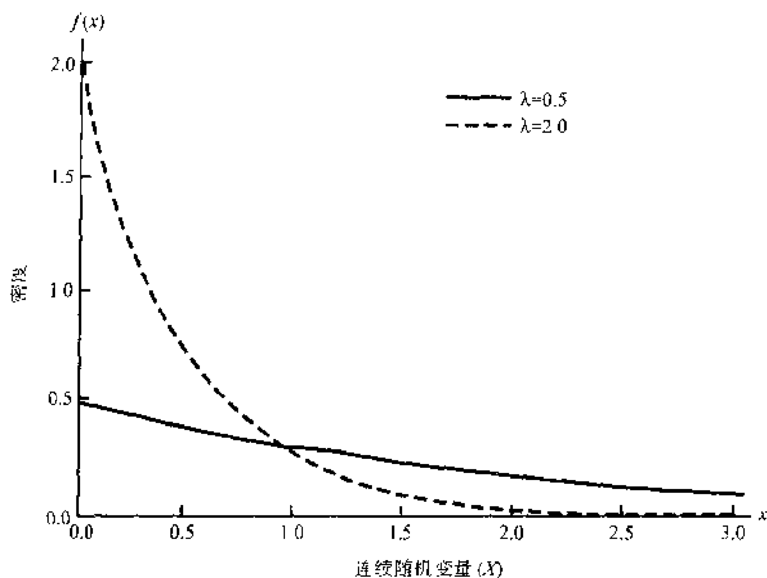


图 12-15

关系:如果一个事件(成功)在某一时间(或空间)单位内由 Poisson 过程(见 11.26 节)随机产生,则连续随机变量-事件相继发生的间隔时间(或空间)-是指数分布的. 而且,如果 λ 是给定单位内的事件平均发生率,则 λ 的倒数,即$\frac{1}{\lambda}$,是事件相继发生的平均间隔时间(或距离).

例 12.15 已知一家医院的急诊室在周日下午 6:00 到 10:00 之间每小时平均到达 $\lambda=5$ 个急救病例. 如果离散随机变量(到达个数)有 Poisson 分布,则在这段时间内:(a) 相继两次到达间隔时间的期望,(b) 前次到达的 15 秒内有另一次到达发生的概率,(c) 对连续变量(相继到达的间隔时间),指数分布的标准差是多少?

解 (a) 由题知,到达次数服从 Poisson 分布,且 λ = 每小时有 5 次到达. 因此,连续随机变量(相继到达间隔时间)服从指数分布,且 $\mu=1/\lambda=1/5=0.2$ 小时. 这是到达间隔的期望时间.

(b) 对连续随机变量(相继到达间隔时间),$\lambda=5$ 的指数概率密度函数[公式(12.27)]是

$$f(x)=5e^{-5x},\ 当\ x\geqslant 0\ 时;\quad f(x)=0\quad 当\ x<0\ 时$$

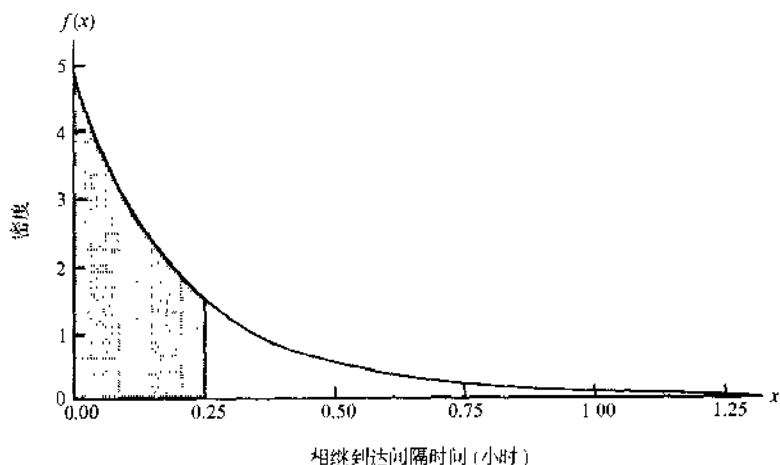


图 12-16

这个指数分布由图 12-16 给出,要求的概率是曲线在区间 $0\leqslant x\leqslant$(15 分钟或 0.25 小时)上的阴影面积. 这个概率值可用公式(12.31)得到:

$$F(x)=P(X\leqslant x)=1-e^{-\lambda x},\ 当\ x\geqslant 0\ 时;\quad F(x)=0\quad 当\ x<0\ 时$$

因此,对 $\lambda=5$ 和 $x=0.25$ 小时,

$$F(0.25)=P(X\leqslant 0.25)=1-e^{-(5\times 0.25)}=1-e^{-1.25}$$

用例 11.22 计算 e^{-x} 的方法,

$$F(0.25)=1-0.2865=0.7135$$

因此,前次到达的 15 秒内有另一次到达发生的概率是 0.7135.

(c) 由公式(12.30),

$$\sigma=1/\lambda=1/5=0.2\text{ 小时}$$

习题解答

正态分布

12.1 对参数为 $\mu=0$ 和 $\sigma^2=4$ 的正态分布,计算 $X=1$ 的 $f(x)$.

解 ☞ 由公式(12.1),

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2}$$

$$f(1)=\frac{1}{\sqrt{4}\times\sqrt{2\pi}}e^{-(1-0)^2/2\times 4}$$

$$=\frac{1}{(2)\times(2.50663)}e^{-1/8}$$

$$=(0.19947)(0.88250)=0.17603$$

标准正态分布

12.2 查表 A.5 求标准正态分布曲线在 $-1.21\leqslant z\leqslant -1.05$ 上的面积.

解 ☞ $-1.21\leqslant z\leqslant 0$ 上的面积是 0.3869, $-1.05\leqslant z\leqslant 0$ 上的面积是 0.3531,从而曲线在 $-1.21\leqslant z\leqslant -1.05$ 上的面积是 $0.3869-0.3531=0.0338$.

12.3 查表 A.5 求标准正态分布曲线在 $0.32\leqslant z\leqslant 1.85$ 上的面积.

解 ☞ $0\leqslant z\leqslant 0.32$ 上的面积是 0.1255, $0\leqslant z\leqslant 1.85$ 上的面积是 0.4878,从而 $0.32\leqslant z\leqslant 1.85$ 上的面积是 $0.4878-0.1255=0.3423$.

12.4 从例 12.6 知,对标准正态分布有 $P(-1.69\leqslant Z\leqslant 0.45)=0.6281$. 则 $P(-1.69<Z<0.45)$ 是多少?

解 ☞ 对于连续型概率分布,上册公式(10.1)表明

$$P(a<X\leqslant b)=P(a\leqslant X\leqslant b)$$
$$=P(a<X<b)=P(a\leqslant X<b)$$

因此,

$$P(-1.69<Z<0.45)=P(-1.69\leqslant Z\leqslant 0.45)=0.6281$$

12.5 上册 7.16 节指出,**经验规则**能粗略的描述总体:对于近似正态分布的总体,≈68%的数据落在区间 $\mu\pm\sigma$ 上,≈95%的数据落在区间 $\mu\pm 2\sigma$ 上,≈100%的数据落在区间 $\mu\pm 3\sigma$ 上.

接着,12.2 节指出,经验规则的精确百分数是 68.3%, 95.4% 和 99.7%. 试用表 A.5 验证这些数值的正确.

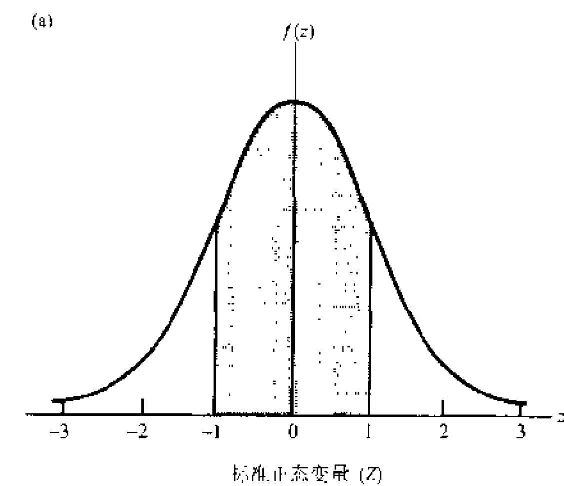
解 ☞ 标准正态分布有均值 $\mu=0$,标准差 $\sigma=1$. 因此,曲线在 $-1\leqslant z\leqslant 1$ 上的面积是分布落在 $\mu-\sigma$ 和 $\mu+\sigma$ 之间的比例(百分数). 这个面积,即图 12-17(a)显示的标准正态分布的阴影区域,是区间 $-1\leqslant z\leqslant 0$ 和 $0\leqslant z\leqslant 1$ 上的相等面积的和. 查表 A.5, $0\leqslant z\leqslant 1$ 上的面积是 0.3413,因此 $\mu\pm\sigma$ 上的分布比例(百分数)是

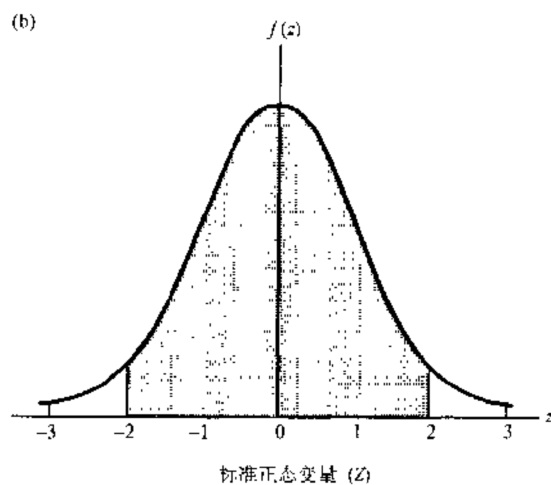
$$2(0.3413) = 0.6826 \text{ 或 } 68.3\%$$

同理，$-2 \leqslant z \leqslant 2$ 上的面积[图 12-17(b)中的阴影]是分布落在 $\mu \pm 2\sigma$ 上的比例(百分数)。查表 A.5，$0 \leqslant z \leqslant 2$ 上的面积是 0.4772，因此 $\mu \pm 2\sigma$ 上的分布比例(百分数)是

$$2(0.4772) = 0.9544 \text{ 或 } 95.4\%$$

最后，$-3 \leqslant z \leqslant 3$ 上的面积[图 12-17(c)中的阴影]是分布落在 $\mu \pm 3\sigma$ 上的比例(百分数)。查表 A.5，$0 \leqslant z \leqslant 3$ 上的面积是 0.4987，因此 $\mu \pm 3\sigma$ 上的分布比例(百分数)是





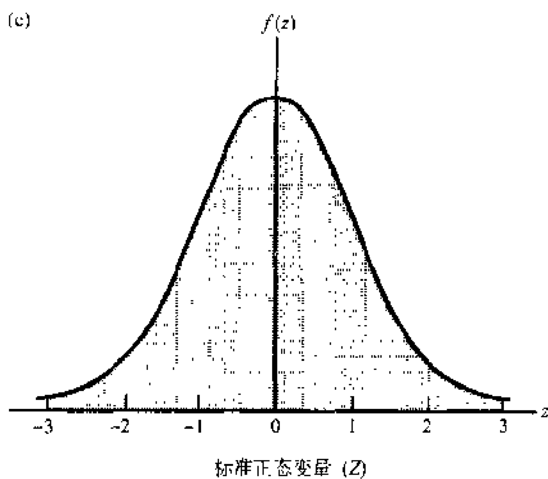


图 12-17

$$2(0.4987)=0.9974 \text{ 或 } 99.7\%$$

由于标准正态分布的区间 $-1\leqslant z\leqslant 1$，$-2\leqslant z\leqslant 2$ 和 $-3\leqslant z\leqslant 3$ 分别表示任何正态分布的 $\mu\pm\sigma$，$\mu\pm 2\sigma$ 和 $\mu\pm 3\sigma$，显然经验规则对所有理论和经验正态分布都成立.

12.6 连续变量 X 服从 $\mu=4.0$ 秒和 $\sigma=0.4$ 秒的正态分布，将下面的 $X=x$ 值变换为 $Z=z$ 值：(a) $x=4.0$ 秒，(b) $x=3.6$ 秒.

解 ☞ (a) 由公式(12.4)，$z=\dfrac{4.0-4.0}{0.4}=0.00$，

(b) $z=\dfrac{3.6-4.0}{0.4}=-1.00$.

12.7 对习题 12.6 中的连续随机变量 X，将下面的 $Z=z$ 值变换回 $X=x$ 值：(a) $z=-2.32$，(b) $z=1.97$.

解 ☞ (a) 由公式(12.12)，$x=4.0$ 秒 $+(-2.32\times 0.4$ 秒$)=4.0$ 秒 -0.928 秒 $=3.1$ 秒，

(b) $x=4.0$ 秒 $+(1.97\times 0.4$ 秒$)=4.0$ 秒 $+0.788$ 秒 $=4.8$ 秒.

12.8 为一家种子公司工作的基因学家，在硬质黏土上培育出一种新品胡萝卜. 对 5000 个这种胡萝卜进行测量，发现：胡萝卜长度 X 服从正态分布，且 $\mu=11.5\text{cm}$，$\sigma=1.15\text{cm}$. X 在区间 $10.0\text{cm}\leqslant x\leqslant 13.0\text{cm}$ 上取值的概率是多少？

解 ☞ 由公式(12.7)，

$$\begin{aligned}&P(10.0\text{cm}\leqslant X\leqslant 13.0\text{cm})\\&=P\left(\frac{10.0\text{cm}-11.5\text{cm}}{1.15\text{cm}}\leqslant Z\leqslant \frac{13.0\text{cm}-11.5\text{cm}}{1.15\text{cm}}\right)\\&=P(-1.30\leqslant Z\leqslant 1.30)\end{aligned}$$

查表 A.5，$0\leqslant z\leqslant 1.30$ 上的面积是 0.4032，这也是 $-1.30\leqslant z\leqslant 0$ 上的面积. 因此

$$\begin{aligned}P(10.0\text{cm}\leqslant X\leqslant 13.0\text{cm})&=P(-1.30\leqslant Z\leqslant 1.30)\\&=2(0.4032)=0.8064\end{aligned}$$

12.9 习题 12.8 的种子公司想在售货清单上标明：新品胡萝卜"能长到 10cm 至 13cm". 为达到这一点，公司要求至少 80% 的胡萝卜能长在 10cm 至 13cm 之间，并且至少 90% 的能长至 10cm 或以上. 公司能否使用这样的措辞？

解 ☞ 由习题 12.8 知

$$P(10.0\text{cm}\leqslant X\leqslant 13.0\text{cm})=0.8064$$

于是，第一个要求能够达到：80.6% 的胡萝卜长在区间 $10.0\text{cm}\leqslant x\leqslant 13.0\text{cm}$ 内.

要知第二个要求能否达到，必须求出

$$P(X\geqslant 10.0\text{cm})=P\left(Z\geqslant \frac{10.0\text{cm}-11.5\text{cm}}{1.15\text{cm}}\right)=P(Z\geqslant -1.30)$$

这个概率由图 12-18 中阴影区域表示.

由于

$$P(Z\geqslant -1.30)=P(-1.30<Z<\infty)=P(-1.30\leqslant Z\leqslant 0)+P(0\leqslant Z<\infty)$$

且

$$P(0\leqslant Z<\infty)=0.5$$

从习题 12.8 知

$$P(-1.30\leqslant Z\leqslant 0)=0.4032$$

于是

$$P(-1.30\leqslant Z<\infty)=0.5+0.4032=0.9032$$

因此，

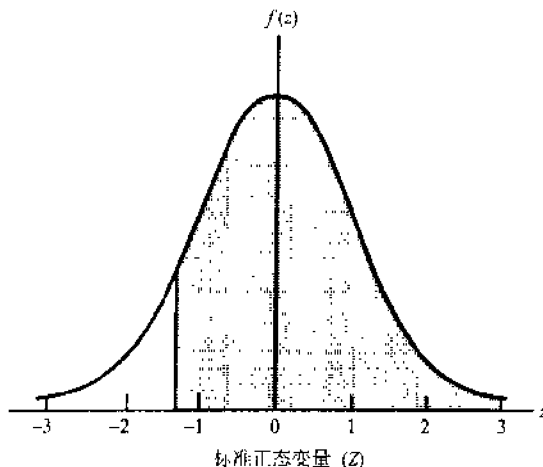


图 12-18

$$P(X \geqslant 10.0\text{cm}) = P(Z \geqslant -1.30) = 0.9032$$

第二个要求:90.3%的胡萝卜生长在10cm或以上能够达到.

12.10　在习题12.8测量的5000个胡萝卜中,有多少个长度为:(a) 10cm和13cm之间,(b) 10cm或更长,(c) 13cm或更短?

解 ☞　(a) 从习题12.8知

$$P(10.0\text{cm} \leqslant X \leqslant 13.0\text{cm}) = 0.8064$$

因此,长度在10cm和13cm之间的胡萝卜个数是

$$0.8064 \times 5000 = 4032$$

(b) 从习题12.9知

$$P(X \geqslant 10.0\text{cm}) = 0.9032$$

因此,长度为10cm或更长的胡萝卜个数是

$$0.9032 \times 5000 = 4516$$

(c) 因为正态分布有对称性,长度为13cm或更短的胡萝卜个数也是4516.

12.11　在习题12.8到12.10的胡萝卜试验中,大于或等于96%的长度是多少?

解 ☞　由于

$$P(X \leqslant x_0) = P(Z \leqslant z_0) = 0.96$$

为解出 x_0,必须首先查表A.5找到 z_0,然后将 z_0 值变换回 x_0. 为找到 z_0,使用以下已知关系

$$\begin{aligned} P(Z \leqslant z_0) = 0.96 &= P(-\infty < Z \leqslant z_0) \\ &= P(-\infty < Z \leqslant 0) + P(0 \leqslant Z \leqslant z_0) \\ &= 0.5 + P(0 \leqslant Z \leqslant z_0) \end{aligned}$$

因此,

$$P(0 \leqslant Z \leqslant z_0) = 0.96 - 0.5 = 0.46$$

表A.5中最接近0.46的面积是0.4599,有 z 值1.75. 于是,由公式(12.12)

$$x_0 = 11.5\text{cm} + (1.75 \times 1.15\text{cm}) = 11.5\text{cm} + 2.0125\text{cm} = 13.5\text{cm}$$

这样,13.5cm的长度大于等于96%的其它长度.

12.12　在习题12.8到12.11的胡萝卜试验中,如果又收获了2个胡萝卜,则这2个都比13.5cm长的概率是多少?

解 ☞　首先，从习题 12.11 知

$$P(X \leqslant 13.5\text{cm}) = 0.96$$

因此，

$$P(X > 13.5\text{cm}) = 1 - 0.96 = 0.04$$

即一个胡萝卜长度超过 13.5cm 的概率是 0.04，如果假设胡萝卜的生长是独立事件，由特殊乘法原理(见上册，9.5 节)知

$$P(A \cap B) = P(A)P(B)$$

如果用 X_1 和 X_2 表示 2 个胡萝卜的长度，则

$$P[(X_1 > 13.5\text{cm}) \cap (X_2 > 13.5\text{cm})] = (0.04) \times (0.04) = 0.0016$$

12.13　一个致力于研究某大城市的刑事诉讼系统的社会学家发现，在过去的 5 年里，被拘捕者从拘捕到审判的等待时间是一个正态分布变量 X，并且 $\mu = 210$ 天，$\sigma = 20$ 天. 被拘捕者中有百分之多少从拘捕到审判的等待时间在 160 天和 190 天之间?

解 ☞　由公式(12.7)，

$$P(160 \leqslant X \leqslant 190) = P\left(\frac{160 - 210}{20} \leqslant Z \leqslant \frac{190 - 210}{20}\right)$$

$$= P(-2.50 \leqslant Z \leqslant -1.00)$$

因为

$$P(-2.50 \leqslant Z \leqslant -1.00) = P(-2.50 \leqslant Z \leqslant 0) - P(-1.00 \leqslant Z \leqslant 0)$$

查表 A.5，

$$P(-2.50 \leqslant Z \leqslant 0) - P(-1.00 \leqslant Z \leqslant 0) = 0.4938 - 0.3413 = 0.1525$$

因此，有 15.25% 的被拘捕者在 160 天至 190 天之后接受审判.

12.14　对某个给定的大总体(例如，美国所有年龄为 12 岁的小孩)，用 Stanford-Binet **智商标准**测得的**智商得分(IQ 得分)**服从 $\mu = 100.0$ 和 $\sigma = 16.0$ 的正态分布，对这个 IQ 分布，第 67 个百分位数的 IQ 得分($X = x$)是多少?

解 ☞　通常用正态分布方法解决 IQ 分布问题，即使目前使用的 IQ 得分总是整数——IQ 变量 X 总是离散变量. 这样处理是将离散型测度变量视为"连续的"，即假定基本的假想变量(即智商)是连续的(见上册，习题 5.9).

要用正态分布求第 67 个百分位数($Q_{67/100} = P_{67}$)，必须首先得到 $x_?$ 值，使得在它之下有 67% 的数据(见上册，6.13 节). 用习题 12.11 中的方法，求解下面公式中的 $x_?$

$$P(X < x_?) = P(X \leqslant x_?) = P\left(Z \leqslant \frac{x_? - 100.0}{16.0}\right) = P(Z \leqslant z_?) = 0.67$$

这里

$$P(Z \leqslant z_?) = 0.67 = P(-\infty < Z \leqslant z_?) = 0.5 + P(0 \leqslant Z \leqslant z_?)$$

于是，

$$P(0 \leqslant Z \leqslant z_?) = 0.67 - 0.5 = 0.17$$

表 A.5 中最接近 0.17 的面积恰好是 0.1700，从而 $z_? = 0.44$.

其次，由于

$$z_? = 0.44 = \frac{x_? - 100.0}{16.0}$$

解出 $x_?$

$$x_? = (16.0 \times 0.44) + 100.0 = 7.04 + 100.0 = 107.04 \text{ 或 } 107.0$$

于是，这个 IQ 分布的第 67 百分位数是 107.0.

12.15　对习题 12.14 中的 IQ 分布：(a) 第 33 个百分位数，(b) 第 2 个十分位数的 IQ 得分各

是多少?

解 ☞ (a) 第 33 个百分位数($Q_{33/100}=P_{33}$)是在它之下有 33% 数据的 x 值. 因此,需解以下方程得到 $x_?$:

$$P(X<x_?)=P(X\leqslant x_?)=P\left(Z\leqslant\frac{x_?-100.0}{16.0}\right)=P(Z\leqslant z_?)=0.33$$

这里

$$P(Z\leqslant z_?)=0.33=0.5-P(z_?\leqslant Z\leqslant 0)$$

于是,

$$P(z_?\leqslant Z\leqslant 0)=0.5-0.33=0.17$$

因此,$z_?$ 是习题 12.14 中 $z_?$ 值的负数形式.

$$z_?=-0.44$$

从而

$$z_?=-0.44=\frac{x_?-100.0}{16.0}$$

并且

$$x_?=[16.0\times(-0.44)]+100.0=-7.04+100.0=92.96\text{ 或 }93.0$$

于是,这个 IQ 分布的第 33 个百分位数是 93.0.

(b) 第 2 个十分位数($Q_{2/10}=D_2$)是在它之下有 20% 的数据的 x 值. 因此,需解以下方程得到 $x_?$:

$$P(X<x_?)=P(X\leqslant x_?)=P\left(Z\leqslant\frac{x_?-100.0}{16.0}\right)=P(Z\leqslant z_?)=0.20$$

这里

$$P(Z\leqslant z_?)=0.20=0.5-P(z_?\leqslant Z\leqslant 0)$$

于是,

$$P(z_?\leqslant Z\leqslant 0)=0.5-0.20=0.30$$

表 A.5 中最接近 0.30 的面积是 0.2996,从而 $z_?=-0.84$. 因此,

$$z_?=-0.84=\frac{x_?-100.0}{16.0}$$

并且

$$x_?=[16.0\times(-0.84)]+100.0=-13.44+100.0=86.56\text{ 或 }86.6$$

于是,这个 IQ 分布的第 2 个十分位数是 86.6.

12.16 对习题 12.14 和 12.15 的 IQ 分布,从总体中随机选出一人,他有以下 IQ 得分的概率是多少:(a) 140,(b) 恰是第 67 个百分位数(从习题 12.14 知是 107.0)?

解 ☞ 如果将这个 IQ 分布看做"连续的"(见习题 12.14),则对于所有的连续型分布[见例 12.1(b)],有

$$P(X=x)=0$$

于是

$$P(X=140)=P(X=107.0)=0$$

如果把每一个精确值看做是所使用的各个测量区间(见上册,2.10 节)的中点,则对 $P(X=140)$ 可以计算 $P(139.5\leqslant X\leqslant 140.5)$,对 $P(X=107.0)$ 可以计算 $P(106.95\leqslant X\leqslant 107.05)$.

(a) 由公式(12.7),

$$P(139.5 \leqslant X \leqslant 140.5) = P\left(\frac{139.5-100.0}{16.0} \leqslant Z \leqslant \frac{140.5-100.0}{16.0}\right)$$
$$= P(2.47 \leqslant Z \leqslant 2.53)$$

查表 A.5 知，$0 \leqslant z \leqslant 2.47$ 上的面积是 0.4932，$0 \leqslant z \leqslant 2.53$ 上的面积是 0.4943. 于是，

$$P(139.5 \leqslant X \leqslant 140.5) = P(2.47 \leqslant Z \leqslant 2.53) = 0.4943 - 0.4932 = 0.0011$$

(b) $$P(106.95 \leqslant X \leqslant 107.05) = P\left(\frac{106.95-100.0}{16.0} \leqslant Z \leqslant \frac{107.05-100.0}{16.0}\right)$$
$$= P(0.43 \leqslant Z \leqslant 0.44)$$

查表 A.5 知，$0 \leqslant z \leqslant 0.43$ 上的面积是 0.1664，$0 \leqslant z \leqslant 0.44$ 上的面积是 0.1700. 于是，

$$P(106.95 \leqslant X \leqslant 107.05) = P(0.43 \leqslant Z \leqslant 0.44) = 0.1700 - 0.1664 = 0.0036$$

单尾和双尾概率

12.17 如果正态分布变量 X 有 $\mu = 14.5$ 和 $\sigma = 2.1$，并且 $\alpha = 0.01$，则 z_α 和 x_α 各是多少？

解 ☞ 由公式(12.9)

$$P(Z > z_\alpha) = \alpha = 0.01$$

因此，为求 z_α，必须查表 A.5 找到最接近 $0.5 - 0.01 = 0.49$ 面积的 z 值：

$$P(0 \leqslant Z \leqslant z_\alpha) = 0.49$$

易得 $z_\alpha = 2.33$.

由公式(12.12)，

$$x_\alpha = 14.5 + (2.33 \times 2.1) = 14.5 + 4.893 = 19.393 \text{ 或 } 19.4$$

12.18 对习题 12.17 中的正态分布变量，$\mu = 14.5$，$\sigma = 2.1$，如果 $\alpha = 0.02$，$-z_\alpha$ 和 $x_{1-\alpha}$ 各是多少？

解 ☞ 由公式(12.11)知

$$P(Z < -z_\alpha) = \alpha = 0.02$$

因此，为求 $-z_\alpha$，必须查表 A.5 找到最接近 $0.5 - 0.02 = 0.48$ 的面积的 z 值：

$$P(-z_\alpha \leqslant Z \leqslant 0) = 0.48$$

这个面积是 0.4798，从而 $-z_\alpha = -2.05$. 将该值代入公式(12.12)，

$$x_{1-\alpha} = 14.5 + (-2.05 \times 2.1) = 14.5 - 4.305 = 10.195 \text{ 或 } 10.2$$

12.19 如果 X 是正态分布变量，且 $\mu = 1.83$，$\sigma = 0.15$，如果 $\alpha = 0.01$，$z_{\alpha/2}$，$-z_{\alpha/2}$，$x_{\alpha/2}$ 和 $x_{1-\alpha}$ 各是多少？

解 ☞ 由公式(12.16)知

$$P(Z > z_{\alpha/2}) = \alpha/2 = 0.01/2 = 0.005$$

因此，为求 z_α，必须查表 A.5 找到最接近 $0.5 - 0.005 = 0.495$ 面积的 z 值：

$$P(0 \leqslant Z \leqslant z_{\alpha/2}) = 0.495$$

存在两个相等的最接近面积：0.4949，此时 $z = 2.57$ 和 0.4951，此时 $z = 2.58$. 因此，

$$z_{\alpha/2} = \frac{2.57 + 2.58}{2} = 2.575$$

因为正态分布有对称性，知

$$-z_{\alpha/2} = -2.575$$

由公式(12.12)，

$$x_{\alpha/2} = 1.83 + (2.575 \times 0.15) = 1.83 + 0.38625 = 2.21625 \text{ 或 } 2.22$$

$$x_{1-\alpha/2} = 1.83 + (-2.575 \times 0.15) = 1.83 - 0.38625 = 1.44375 \text{ 或 } 1.44$$

二项分布的正态近似

12.20 在例 12.10(a),对于二项变量:掷 14 次硬币出现正面的次数,已解得出现 9 次正面的概率是 $P(X=9)=0.1222$. 对这个变量,试用二项分布的正态近似计算 $P(X=9)$.

解 把 9 看做是从 8.5 到 9.5 的一个测量值,问题是求图 12-12($\mu = np = 7$, $\sigma = \sqrt{npq} = \sqrt{3.5} = 1.8708$)中区间 $8.5 \leqslant x \leqslant 9.5$ 上正态曲线下方的面积. 由例 12.10(b)的公式(12.7),

$$P(8.5 \leqslant X \leqslant 9.5) = P\left(\frac{8.5 - np}{\sqrt{npq}} \leqslant Z \leqslant \frac{9.5 - np}{\sqrt{npq}}\right)$$

$$= P\left(\frac{8.5 - 7}{1.8708} \leqslant Z \leqslant \frac{9.5 - 7}{1.8708}\right) = P(0.80 \leqslant Z \leqslant 1.34)$$

查表 A.5,

$$P(0 \leqslant Z \leqslant 0.80) = 0.2881 \quad P(0 \leqslant Z \leqslant 1.34) = 0.4099$$

因此,

$$P(0.80 \leqslant Z \leqslant 1.34) = 0.4099 - 0.2881 = 0.1218$$

比较二项真值和正态近似值,可以看到它们在前两位小数上相同:

$$P(X = 9) = 0.1222 \approx P(0.80 \leqslant Z \leqslant 1.34) = 0.1218$$

12.21 对于二项变量:掷 14 次硬币出现正面的次数,分别用二项分布和正态近似方法(见例 12.10)计算至多出现 6 次正面的概率.

解 计算二项概率真值 $P(X \leqslant 6)$,需要用到例 12.10(a)中二项式展开的有关项:

$$P(X = 0) = q^{14} = \left(\frac{1}{2}\right)^{14} = 0.000061$$

$$P(X = 1) = 14q^{13}p = 14\left(\frac{1}{2}\right)^{13}\left(\frac{1}{2}\right) = 0.000854$$

$$P(X = 2) = 91q^{12}p^2 = 91\left(\frac{1}{2}\right)^{12}\left(\frac{1}{2}\right)^2 = 0.005554$$

$$P(X = 3) = 364q^{11}p^3 = 364\left(\frac{1}{2}\right)^{11}\left(\frac{1}{2}\right)^3 = 0.022217$$

$$P(X = 4) = 1001q^{10}p^4 = 1001\left(\frac{1}{2}\right)^{10}\left(\frac{1}{2}\right)^4 = 0.061096$$

$$P(X = 5) = 2002q^9p^5 = 2002\left(\frac{1}{2}\right)^9\left(\frac{1}{2}\right)^5 = 0.122192$$

$$P(X = 6) = 3003q^8p^6 = 3003\left(\frac{1}{2}\right)^8\left(\frac{1}{2}\right)^6 = 0.183289$$

因此,

$$P(X \leqslant 6) = 0.000061 + 0.000854 + 0.005554 + 0.022217 + 0.061096 + 0.122192 + 0.183289$$
$$= 0.395263 \text{ 或 } 0.3953$$

正态近似解只须求出正态曲线在以下两个区间:$-0.5 \leqslant x \leqslant 6.5$ 和 $-\infty < x \leqslant 6.5$ 的任意一个上的面积即可. 第一个区间,$-0.5 \leqslant x \leqslant 6.5$,表示端点经过连续性修正的二项分布的精确区间(0 到 6). 第二个区间,$-\infty < x \leqslant 6.5$,在应用中非常典型,经过连续性修正从上边界 6.5 延伸到正态分布的下极限 $-\infty$.

对 $-0.5 \leqslant x \leqslant 6.5$ 求解,用 12.12 节的标准正态分布,有 $\mu = np = 7$, $\sigma = \sqrt{npq} = \sqrt{3.5} = 1.8708$:

$$P(-0.5 \leqslant X \leqslant 6.5) = P\left(\frac{-0.5 - np}{\sqrt{npq}} \leqslant Z \leqslant \frac{6.5 - np}{\sqrt{npq}}\right)$$

$$= P\left(\frac{-0.5-7}{1.8708} \leqslant Z \leqslant \frac{6.5-7}{1.8708}\right)$$

$$= P(-4.01 \leqslant Z \leqslant -0.27)$$

查表A.5

$$P(-4.01 \leqslant Z \leqslant 0) = 0.5000 \quad P(-0.27 \leqslant Z \leqslant 0) = 0.1064$$

因此,

$$P(-4.01 \leqslant Z \leqslant -0.27) = 0.5000 - 0.1064 = 0.3936$$

对 $-\infty < x \leqslant 6.5$ 求解,从上面结果知

$$P(-\infty < X \leqslant 6.5) = P(-\infty < X \leqslant -0.27)$$

并且

$$P(-\infty < X \leqslant -0.27) = 0.5000 - 0.1064 = 0.3936$$

用两个区间得到的结果是相同的,而且几乎总是这样. 由于计算端点处($-\infty$或∞)的正态分布更容易,故这个解法更为常用.

比较二项真值和正态近似解,可以看到在前二位小数上它们是相同的:

$$P(X \leqslant 6) = 0.3953 \approx P(-\infty < Z \leqslant -0.27) = 0.3936$$

12.22 在某城市,注册选民的48%是共和党人. 电话随机抽样200个注册选民,并询问他们是否支持兴建一座新机场. 使用正态近似方法,电话询问的选民中至多有100个共和党人的概率是多少?

解 假设这是一个二项试验,则要求的是二项分布中的 $P(X<100)$,这里,$\mu = np = 200 \times 0.48 = 96$, $\sigma^2 = npq = 200 \times 0.48 \times 0.52 = 49.92$, $\sigma = \sqrt{npq} = \sqrt{49.92} = 7.0654$. 因为 np 和 nq 都大于5($np = 96$, $nq = 104$),故可以使用正态近似解法(见12.12节).

由于二项变量的所有取值均为整数,因此

$$P(X < 100) = P(X \leqslant 99)$$

进行连续性修正[见12.12(b)],正态近似值是正态曲线($\mu = 96$, $\sigma^2 = 49.92$, $\sigma = 7.0654$)在区间 $-\infty < x \leqslant 99.5$ 上的面积:

$$P(-\infty < X \leqslant 99.5) = P\left(-\infty < Z \leqslant \frac{99.5 - np}{\sqrt{npq}}\right)$$

$$= P\left(-\infty < Z \leqslant \frac{99.5 - 96}{7.0654}\right) = P(-\infty < Z \leqslant 0.50)$$

查表A.5,

$$P(0 \leqslant Z \leqslant 0.50) = 0.1915$$

于是,

$$P(-\infty < Z \leqslant 0.50) = 0.5000 + 0.1915 = 0.6915$$

因此,

$$P(X < 100) \approx 0.6915$$

12.23 某制造商接收了批量为100000的一批保险丝. 他的抽样计划(见11.11节)是从每一批随机抽取800根保险丝作为样本并检验,如果样本含有10根或更少的不合格品则接受相应批. 假设批量的不合格品率为 $p = 0.01$,用正态近似方法计算该批的接受概率(P_a)?

解 假设这是一个二项试验,则要求的是二项分布中的 $P_a = P(X \leqslant 10)$,这里,$\mu = np = 800 \times 0.01 = 8$, $\sigma^2 = npq = 800 \times 0.01 \times 0.99 = 7.92$, $\sigma = \sqrt{npq} = \sqrt{7.92} = 2.8142$. 因为 np 和 nq 都大于5($np = 8$, $nq = 792$),故可以使用正态近似解法(见12.12节).

进行连续性修正[见 12.12(b)],正态近似值是正态曲线($\mu=8$, $\sigma^2=7.92$, $\sigma=2.8142$)在区间 $-\infty<x\leqslant10.5$ 上的面积:

$$P(-\infty<X\leqslant10.5)=P\left(-\infty<Z\leqslant\frac{10.5-np}{\sqrt{npq}}\right)$$

$$=P\left(-\infty<Z\leqslant\frac{10.5-8}{2.8142}\right)=P(-\infty<Z\leqslant0.89)$$

查表 A.5,

$$P(0\leqslant Z\leqslant0.89)=0.3133$$

于是,

$$P(-\infty<Z\leqslant0.89)=0.5000+0.3133=0.8133$$

因此,

$$P_a=P(X\leqslant10)\approx0.8133$$

Poisson 分布的正态近似

12.24 例 12.11 中的电缆制造商已经确知,对 4 米为一单位的电缆,每 4 米有 $\lambda=4.0$ 个缺陷. 分别用 Poisson 分布和正态近似方法计算在 $t=4.5$ 单位的电缆上有 17,18 或 19 个缺陷的概率.

解 ☞ 计算 Poisson 概率真值 $P(17\leqslant X\leqslant19)$,需要用到公式(11.26),

$$f(x)=\frac{\mu^x e^{-\mu}}{x!}$$

找到 Poisson 分布的 $f(17)$, $f(18)$和 $f(19)$,这里 $\mu=\sigma^2=\lambda t=(4.0)(4.5)=18$, $\sigma=\sqrt{\lambda t}=\sqrt{18}=4.2426$:

$$f(17)=\frac{(18)^{17}e^{-18}}{17!}=0.0936$$

$$f(18)=\frac{(18)^{18}e^{-18}}{18!}=0.0936$$

$$f(19)=\frac{(18)^{19}e^{-18}}{19!}=0.0887$$

由例 12.10(a)知,对诸如这样的离散变量,

$$P(17\leqslant X\leqslant19)=P(X=17)+P(X=18)+P(X=19)$$

$$=f(17)+f(18)+f(19)=0.0936+0.0936+0.0887=0.2759$$

由于 $\lambda t\geqslant10$,故可以使用 12.13 节的正态近似方法. 并且进行连续性修正后,只须计算 $P(16.5\leqslant X\leqslant19.5)$,此时,正态分布有 $\mu=\sigma^2=\lambda t=18$, $\sigma=\sqrt{18}=4.2426$,于是

$$P(16.5\leqslant X\leqslant19.5)=P\left(\frac{16.5-\lambda t}{\sqrt{\lambda t}}\leqslant Z\leqslant\frac{19.5-\lambda t}{\sqrt{\lambda t}}\right)$$

$$=P\left(\frac{16.5-18}{4.2426}\leqslant Z\leqslant\frac{19.5-18}{4.2426}\right)=P(-0.35\leqslant Z\leqslant0.35)$$

查表 A.5 知

$$P(-0.35\leqslant Z\leqslant0)=P(0\leqslant Z\leqslant0.35)=0.1368$$

因此

$$P(-0.35\leqslant Z\leqslant0.35)=2(0.1368)=0.2736$$

比较正态近似值和 Poisson 真值,可以看到在前二位小数上它们是相同的:

$$P(17\leqslant X\leqslant19)=0.2759\approx P(-0.35\leqslant Z\leqslant0.35)=0.2736$$

这里得到的正态近似值比在例 12.11 中得到的值更接近于 Poisson 真值,因为在这个问题中 λt 增大了三倍(18 对 6). 从 12.13 节知,λt 越大,正态近似值越接近真值.

12.25 对于例 11.25 的 Geiger 计数器试验,物理学家已确知:以 10 秒为一单位,每 10 秒有 $\lambda=5.5$ 个粒子. 使用正态近似方法计算在 2 分钟内记录的粒子多于 73 个的概率.

解 ☞ 因为所有 Poisson 变量的取值均是整数,

$$P(X>73)=P(X\geqslant 74)=P(74\leqslant X)$$

于是,进行连续性修正,正态近似值是正态曲线($\mu=\sigma^2=\lambda t=5.5\times 12=66$, $\sigma=\sqrt{\lambda t}=8.1240$)在区间 $73.5\leqslant x<\infty$ 上的面积:

$$P(73.5\leqslant X<\infty)=0.5000-P\left(0\leqslant Z\leqslant\frac{73.5-\lambda t}{\sqrt{\lambda t}}\right)$$

$$=0.5000-P\left(0\leqslant Z\leqslant\frac{73.5-66}{8.1240}\right)=0.5000-P(0\leqslant Z\leqslant 0.92)$$

查表 A.5 知

$$P(73.5\leqslant Z<\infty)=0.5000-0.3212=0.1788$$

因此,

$$P(X>73)\approx 0.1788$$

均匀概率分布

12.26 对掷骰子试验,用均匀概率分布[见例 12.12(f)]计算 $P(X<4)$.

解 ☞ 将 $k=6$ 代入公式(12.19),

$$f(x)=\frac{1}{k}=\frac{1}{6}$$

并且

$$P(X<4)=P(X=1)+P(X=2)+P(X=3)$$

$$=f(1)+f(2)+f(3)=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{3}{6}=0.5$$

12.27 对于例 12.13 中的连续随机变量 X,只在区间 $1\leqslant x\leqslant 6$ 上取值. 如果 X 有连续型均匀概率分布,计算 $P(X<4)$.

解 ☞ 由于在区间 $1\leqslant x\leqslant 6$ 以外 $f(x)=0$,并且由于这是连续型概率分布,

$$P(X<4)=P(X\leqslant 4)=P(1\leqslant X\leqslant 4)$$

因此,用来自例 12.13(f)的公式,

$$P(c\leqslant X\leqslant d)=\frac{d-c}{b-a}$$

其中 $d=4$, $c=1$, $b=6$, $a=1$:

$$P(X<4)=\frac{d-c}{b-a}=\frac{4-1}{6-1}=\frac{3}{5}=0.6$$

12.28 在关于鸟类导航的研究中,研究者将幼鸟置于一个巨大的圆形笼子中央,在不同的试验条件下观察幼鸟从中央飞起的方向. 笼子的周长从正北开始按顺时针方向以角度来测量,正北既为 0°又为 360°. 对于每次飞起,研究者以正北开始的顺时针角度值记录幼鸟飞起方向. 在每次试验中,研究者不提供任何实物暗示,因此可以假定飞起方向是随机决定的. 这样,连续随机变量——从正北开始的角度——服从区间 $0\leqslant x\leqslant 360$ 上的均匀分布. 假设这是对的,画出这个分布的图形,标出密度函数 $f(x)$ 和 $\mu\pm\sigma$,并计算一次飞起落在从正北开始的 10°到 50°之间的概率.

解 ☞ 从 12.15 节知

$$f(x)=\frac{1}{b-a},\text{当 } a\leqslant x\leqslant b \text{ 时};f(x)=0,\text{其它}$$

$$\mu = \frac{b+a}{2}$$

$$\sigma = \sqrt{\frac{(b-a)^2}{12}}$$

本问题中,

$$f(x) = \frac{1}{360-0} = \frac{1}{360}, \text{当 } 0 \leqslant x \leqslant 360 \text{ 时}; f(x) = 0, \text{其它}$$

$$\mu = \frac{360+0}{2} = 180$$

$$\sigma = \sqrt{\frac{(360-0)^2}{12}} = \sqrt{10800} = 103.92$$

这些值标注在图 12-19 所示的均匀分布的图形上.

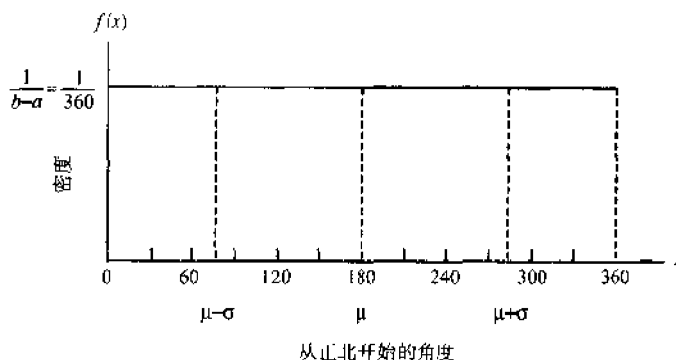


图 12-19

从例 12.13(f)知,X 在定义区间 $a \leqslant x \leqslant b$ 内的某区间 $c \leqslant x \leqslant d$ 上取值的概率是$\frac{d-c}{b-a}$. 因此,对 $a=0$, $b=360$, $c=10$, $d=50$,

$$P(10 \leqslant X \leqslant 50) = \frac{50-10}{360-0} = 0.11$$

12.29 假设你要参加在 11 层召开的一个会议,并且在会议开始前 5 分钟,你刚好到达 10 层的电梯口. 已知在这幢大楼的任意一层等待电梯花费的时间在 0 到 10 分钟之间变化,而且这个连续随机变量(等待时间)服从区间 $0 \leqslant x \leqslant 10$ 上的均匀分布. 如果电梯运行一层所用时间为 10 秒,从 11 层电梯口到达会场所用时间为 20 秒,则你能准时到达会场的概率是多少?

解 ☞　在区间中加入固定花费时间 30 秒,得到一个新变量(到会时间),它服从区间 $0.5 \leqslant x \leqslant 10.5$上的均匀分布. 对该变量,需要计算 $P(0.5 \leqslant X \leqslant 5)$. 从例 12.13(f)知,$X$ 在定义域 $0.5 \leqslant x \leqslant 10.5$ 内的区间 $0.5 \leqslant x \leqslant 5$ 上取值的概率值是

$$P(0.5 \leqslant X \leqslant 5) = \frac{d-c}{b-a} = \frac{5-0.5}{10.5-0.5} = \frac{4.5}{10} = 0.45$$

12.30 一个企业心理学家已确知,工人在自动流水线上完成一项任务所需时间在 9 到 15 分钟之间. 如果连续随机变量——完成任务所需时间——服从区间 $9 \leqslant x \leqslant 15$ 上的均匀分布,则对该分布计算:(a) $f(x)$, μ, σ,和参数,(b) $P(X<13)$, (c) $P(4<X<7)$.

解 ☞　(a) 由公式(12.23),(12.24)和(12.26),

$$f(x) = \frac{1}{b-a} = \frac{1}{15-9} = \frac{1}{6}, \text{当 } 9 \leqslant x \leqslant 15 \text{ 时}; f(x) = 0, \text{其它}$$

$$\mu = \frac{b+a}{2} = \frac{15+9}{2} = 12$$

$$\sigma = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(15-9)^2}{12}} = \sqrt{3} = 1.73$$

参数 $a=9$，$b=15$.

(b) 用例 12.13(f)中公式，

$$P(X < 13) = P(9 \leqslant X \leqslant 13) = \frac{d-c}{b-a} = \frac{13-9}{15-9} = 0.67$$

(c) $P(4<X<7)=0$，因为在区间 $9\leqslant x\leqslant 15$ 以外 $f(x)=0$.

指数分布

12.31 对于例 12.15 的医院，周日下午 6:00 到 10:00 之间，平均每小时有 5 例急诊到达急诊室. 图 12-16 给出随机变量(到达间隔时间)的指数分布，则相继两次到达间隔时间超过 10 分钟的概率是多少?

解 ☞ 同任何连续型概率分布一样，指数分布曲线(从 0 到∞)下方的总面积(概率)是 1.00 [见例 12.1(f)]. 因此，

$$P(X > x) = 1 - P(X \leqslant x)$$

由公式(12.31)知

$$P(X \leqslant x) = F(x) = 1 - e^{-\lambda x}$$

因此

$$P(X > x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$$

于是，对 $x=(10$ 分钟或 0.1667 小时)和 $\lambda=5$，

$$P(X > 0.1667) = e^{-(5\times 0.1667)} = e^{-0.8335} = 0.4345$$

因此，相继两次到达间隔时间超过 10 分钟的概率是 0.4345.

12.32 对例 12.31 的急诊室，如果离散随机变量(到达次数)服从 Poisson 分布，则从 6:00 到 10:00 之间的任何整点起，下次到达所需时间在 15 到 25 分钟之间的概率是多少?

解 ☞ 这个问题与例 12.15(b)和习题 12.31 的问题类似，只是现在感兴趣的是给定时间间隔内下次到达所需时间，而不是相继两次到达间隔时间. 我们说指数分布的随机变量是**无记忆的**，即未来事件的概率与诸如过程持续的时间或一个事件多久发生等因素无关. 因此，在这个问题中，对于整点是前一个到达点还是任意间隔时间，下次到达所需时间的概率是相同的.

因此，连续随机变量，下次到达所需时间的 $\lambda=5$ 的指数概率分布仍然是[见例 12.15(b)]

$$f(x) = 5e^{-5x},\ x \geqslant 0;\ f(x) = 0,\ x < 0$$

这个指数分布如图 12-20 所示，所求概率是曲线在区间(15 分钟或 0.25 小时)$\leqslant x \leqslant$(25 分钟或 0.4167小时)上的阴影面积. 概率值可以用上册公式(10.9)得到. 即

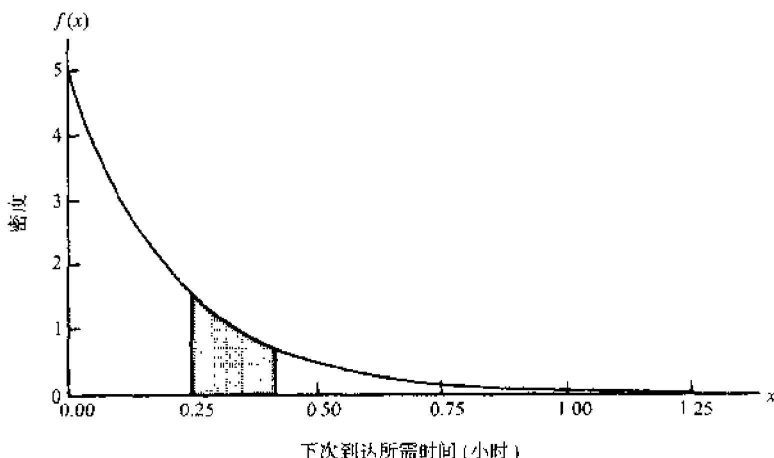


图 12-20

$$P(0.25 \leqslant X \leqslant 0.4167) = F(0.4167) - F(0.25)$$
$$= [1 - e^{-(5\times 0.4167)}] - [1 - e^{-(5\times 0.25)}] = [1 - e^{-2.0835}] - [1 - e^{-1.25}]$$
$$= (1 - 0.1245) - (1 - 0.2865) = 0.1620$$

因此,在6:00到10:00之间的任何间隔内,下次到达所需时间在15到25分钟之间的概率是0.1620.

12.33 在正常使用下,一种电视显像管的**失效时间**是7.3年.如果失效时间服从Poisson分布,则这样的一个显像管从开始使用在4年内失效的概率是多少?

解 ☞　指数分布经常用于对一个系统的失效时间(或**生存时间**或**寿命**)进行建模.这种应用要求系统有指数分布的无记忆性(见习题12.32).这里,要求有假设,系统的未来寿命与系统的运行时间独立.尽管在现实世界中这种要求几乎不能达到,但是,如果平均失效时间很长,如本问题,则指数分布是一个合理的模型.

假设电视显像管的失效时间服从指数分布,可知

$$\mu = 1/\lambda = 7.3\text{年}$$

从而

$$\lambda = 1/7.3 = 0.1370$$

因此,由公式(12.31),

$$F(4) = P(X \leqslant 4) = 1 - e^{-(0.1370\times 4)} = 1 - e^{-0.5480} = 1 - 0.5781 = 0.4219$$

于是,这样一个显像管从开始使用在4年内失效的概率是0.4219.

12.34 对习题12.33的电视显像管:独立地用在不同电视机中的4个显像管,从开始使用在4年内失效的概率是多少?

解 ☞　因为每个显像管的失效是独立事件,所以可以使用特殊乘法原理[上册,公式(9.9)]的推广形式解决这个问题.从习题12.33知

$$P(X \leqslant 4) = 0.4219$$

如果用X_1, X_2, X_3和X_4表示四个失效时间,则

$$P[(X_1 \leqslant 4) \cap (X_2 \leqslant 4) \cap (X_3 \leqslant 4) \cap (X_4 \leqslant 4)] = (0.4219)^4 = 0.0317$$

因此,4个这样的显像管从开始使用在4年内失效的概率是0.0317.

12.35 对习题12.33和12.34的电视显像管,已知一个显像管在h年或更短时间内失效的概率是0.8,则h是多少?

解 ☞　从习题12.33知

$$F(h) = P(X \leqslant h) = 1 - e^{-0.1370h}$$

因此,

$$1 - e^{-0.1370h} = 0.8$$

并且

$$e^{-0.1370h} = 0.2$$

从上册1.10节知

$$\text{如果} \log_e n = x, \text{则} n = e^x$$

因此,在这里[见上册,习题1.23(b)]

$$\log_e 0.2 = -0.1370h$$
$$-1.6094 = -0.1370h$$

从而,

$$h = \frac{-1.6094}{-0.1370} = 11.7474\text{年}$$

补充习题

正态和标准正态分布

12.36　用公式(12.1),对参数 $\mu=2$ 和 $\sigma^2=1$ 的正态分布计算 $X=1$ 的 $f(x)$.

答案:0.24197

12.37　连续随机变量 X 服从 $\mu=4.0$ 秒和 $\sigma=0.4$ 秒的正态分布,将以下 $X=x$ 值变换为 $Z=z$ 值:(a) $x=5.3$ 秒,(b) $x=3.3$ 秒.

答案:(a) 3.25, (b) −1.75

12.38　对习题 12.37 的连续随机变量 X,将以下 $Z=z$ 值变换回 $X=x$ 值:(a) $z=-0.92$, (b) $z=2.93$.

答案:(a) 3.6 秒, (b) 5.2 秒

12.39　对于习题 12.13 的刑事诉讼研究,随后的 100 个被拘捕者中将有多少个能在 200 天以内审判?

答案:30.85 或 31

12.40　根据习题 12.13 和 12.39 的刑事诉讼的研究结果,反对随后的一个被捕者在 200 天之内审判有多大的胜算?

答案:$P(A)$=被捕者在 200 天之内审判的概率. $\frac{P(A')}{P(A)}=\frac{1-0.3085}{0.3085}=\frac{0.6915}{0.3085}\approx\frac{7}{3}$;反对审判在 200 天之内进行的胜算大约是 7 比 3.

12.41　QUICK-TUNE 公司做广告声称完全翻新一辆汽车只需 \$34.98. 而且,公司保证,如果不能在 30 分钟或更短的时间内完成,翻新将免费. 公司已确知一次翻新所需时间是一个正态分布变量,且 $\mu=23.2\text{min}$, $\sigma=4.17\text{min}$. 如果 QUICK-TUNE 一次翻新的实际成本平均为 \$24.00(包括配件,劳动力,广告费,等等),则在接下来的 200 次翻新中公司的期望利润是多少?

答案:$P(X>30\text{min})=0.5-0.4484=0.0516$;免费翻新的次数 $=0.0516\times200=10.32$ 或 10;免费翻新的成本 $=10\times\$24=\$240$;非免费翻新的利润 $=190\times(\$34.98-\$24.00)=\$2086.20$;总利润 $=\$2086.20-\$240.00=\$1846.20$.

12.42　参加大学心理学课程学习的 380 个学生均有机会在该课程得到 500 分的满分. 学生得分 X 服从正态分布,并且教授将画出学生得分的"等级曲线". 得分和等级之间的关系如下:A,如果得分落在区间 $x\geqslant(\mu+\sigma)$; B,如果得分落在区间 $\mu\leqslant x<(\mu+\sigma)$;$C$,如果得分落在区间 $(\mu-\sigma)\leqslant x<\mu$;$D$,如果得分落在区间 $(\mu-2\sigma)\leqslant x<(\mu-\sigma)$;$F$,如果得分落在区间 $x<(\mu-2\sigma)$. 如果可以证明 A 和 C 的最低得分分别是 448 和 352,则:(a) σ 和 μ 是多少? (b) 多少个学生得到 B?

答案:将这个分布看做是"连续的":(a) $\sigma=48$, $\mu=400$, (b) 129.69 或 130

单尾和双尾概率

12.43　对例 12.7 的地鼠,雄性的体重服从均值为 400g 和标准差为 100g 的正态分布. 如果 $\alpha=0.01$,则 z_α 和 x_α 是多少?

答案:$z_\alpha=2.33$, $x_\alpha=633\text{g}$

12.44　对习题 12.43 的雄性地鼠的体重,如果 $\alpha=0.02$,则 $-z_\alpha$ 和 $x_{1-\alpha}$ 是多少?

答案:$-z_\alpha=-2.05$, $x_{1-\alpha}=195\text{g}$

12.45　对习题 12.43 的雄性地鼠的体重,如果 $\alpha=0.05$,则 $z_{\alpha/2}$, $-z_{\alpha/2}$, $x_{\alpha/2}$ 和 $x_{1-\alpha/2}$ 是多少?

答案:$z_{\alpha/2}=1.96$, $-z_{\alpha/2}=-1.96$, $x_{\alpha/2}=596\text{g}$, $x_{1-\alpha/2}=204\text{g}$

二项分布的正态近似

12.46　对二项变量:掷 800 次硬币出现正面的次数,用正态近似方法计算至少出现 415 次正面的概率.

答案:≈0.1357

12.47　一名皮肤病专家对某种皮肤病采用药物 A 处理,已知治愈率为 64%. 用正态近似方法计算:将药物 A 用于随后的 250 个患这种皮肤病的病人,治愈多于 150 个而少于 175 个的概率.

答案:≈0.8663

12.48　一所药物学校为下一年招生从 130 个地区收到 1000 份申请. 学校计划招生 130 名. 以往经验表明,通常只有 60%的接收者将最终注册入学. 如果学校接收 200 个申请者,则有太多的申请者——131 个或更多——愿意注册入学的概率是多少? 用正态近似方法解答这个问题.

答案:≈0.0643

Poisson 分布的正态近似

12.49 对例 11.25 和习题 12.25 的 Geiger 计数器试验,物理学家计算出 10 秒钟内受激粒子的平均比例是 $\lambda=5.5$ 个. 用正态近似方法,计算在 2 分钟内粒子数为 66 个的概率.

答案:≈ 0.0478

12.50 每小时有 30 辆汽车经过路口. 在某 5 分钟内至少有 6 辆汽车经过路口的概率是多少?

答案:≈ 0.0057

均匀概率分布

12.51 对掷骰子试验,用均匀概率分布计算 $P(5\leqslant X\leqslant 9)$.

答案:0.3333

12.52 连续随机变量 X 只能在区间 $1\leqslant X\leqslant 6$ 上取值. 如果 X 服从连续型均匀分布,则 $P(5\leqslant X\leqslant 9)$是多少?

答案:0.2000

12.53 在习题 12.28 的鸟类飞起方向研究中,研究员假定飞起方向是随机决定的,且随机变量(偏北度数)服从区间 $0\leqslant X\leqslant 360$ 上的均匀分布. 假设他是对的,计算一次飞起将在偏北 300°和 340°之间的概率是多少?

答案:0.1111

12.54 在习题 12.30 中的企业心理学家已经确知,工人在自动流水线上完成一项任务需用 9 到 15 分钟时间. 如果连续随机变量——完成任务所需时间,服从区间 $9\leqslant X\leqslant 15$ 上的均匀分布,则对该分布计算:(a)$P(12<X<14)$,(b)$P(X\geqslant 10)$.

答案:(a)0.3333,(b)0.8333

指数分布

12.55 对 11.26 节的电缆缺陷试验,离散随机变量(缺陷个数)服从 $\lambda=$ 每 4 米 4.0 个缺陷的 Poisson 分布,试问:两个缺陷至少相隔 50cm 的概率是多少?

答案:$e^{-0.5}=0.6065$

12.56 对习题 12.50 经过汽车数的研究,每小时有 30 辆汽车经过路口,如果离散随机变量(经过的汽车数)服从 Poisson 分布,则连续两次经过的间隔时间至少为 5 分钟的概率是多少?

答案:0.0822

12.57 一个鸟类学家沿着一道篱笆走过某种鸟的栖息地,记录下能够看到的鸟巢个数. 在 1 公里(1000 米)上他记录了 30 个鸟巢. 如果离散随机变量(鸟巢个数)服从 Poisson 分布,则两个鸟巢相隔距离至多为 20 米的概率是多少?

答案:0.4512

12.58 对习题 12.57 描述的某种鸟巢个数的研究,从篱笆上的任何一点到下一个鸟巢之间的距离在 10 到 30 米之间的概率是多少?

答案:0.3342

第十三章　抽样分布

13.1　简单随机抽样回顾

在上册第三章中，我们指出统计科学分为两部分：描述性统计学（收集，整理，描述和表示数据）和推断性统计学（从样本信息推断整个总体的信息）.并且指出推断性统计学有4个理论组成部分：概率论，抽样理论，估计理论和假设检验理论.在统计学的这些基础理论中，我们已经详细介绍了描述性统计学（见上册，第二章至第七章）和概率论（见上册，第八章至第十章；和本册的第十一章和第十二章）.现在，继续介绍余下的三个组成部分：抽样理论（本章），估计理论（从第十四章开始）和假设检验理论（从第十六章开始）.

上册第三章简要讨论了抽样理论的一些知识，重点放在抽样方法，称为抽样设计（见3.15节），它是推断性统计学的要求.同时指出，一些推断性方法，即所谓的随机抽样或概率抽样，可用于任何形式的抽样设计（见3.17节）.但是其中最重要的就是简单随机抽样（见3.18节），事实上，初等推断性统计学的大部分方法都是基于这样一个假设，即样本是通过简单随机抽样得到的.

13.2　独立随机变量

上册3.14节讨论了数理统计学和一般统计学的差异.数理统计学是包含公理和定理的一个完整的数学体系.相比较，一般统计学是这个数学体系的非数学解释，是为了满足非统计学家的一般性的实际应用.因此，简单随机抽样既有基本的数学定义，也有许多对该定义的简单的非数学解释（见13.3节）.

简单随机抽样的理论和数学定义的一个重要组成是**独立随机变量**的概念.由概率论（见上册，9.4节）知，两个事件A和B相互独立，如果其中一个发生不影响另一个是否发生的概率.类似地，两个随机变量X和Y相互独立，如果当X取定一个x值时，不影响Y取某个y值的概率.

对离散随机变量X，Y和Z，如果事件$X=x$，$Y=y$和$Z=z$对一切x，y和z相互独立，则变量X，Y和Z相互独立，而且由特殊乘法原理的推广［上册，公式(9.9)］

$$P(X=x,\ Y=y,\ Z=z)=P(X=x)P(Y=y)P(Z=z)$$

对连续随机变量X，Y和Z，如果事件$X=x$，$Y=y$和$Z=z$对一切x，y和z相互独立，则

$$P(X\leqslant x,\ Y\leqslant y,\ Z\leqslant z)=P(X\leqslant x)P(Y\leqslant y)P(Z\leqslant z)$$

13.3　简单随机抽样的数学定义和非数学定义

对总体的**简单随机抽样的数学定义**是用随机变量和概率分布的语言来表述的.回顾10.5和10.6节，概率分布是总体相对频数分布的数学模型.也就是说，如果测量值总体是由随机变量X产生的，则变量X的概率分布就是该总体分布的数学模型.

如果把取自总体的容量为n的样本看做是一个统计试验进行n次的结果，则每次试验的结果是对总体随机变量X的一个观测.因为n次观测的每一次都是随机决定的（见上册，10.1节），所以n次观测对应n个随机变量$X_1,X_2,\cdots,X_n$的集合.如果n个随机变量满足两个条件：(1) n个随机变量相互独立（见13.2节），(2) n个随机变量都有与总体随机变量X相同的概率分布，则n个变量$X_1,X_2,\cdots,X_n$的集合称为一个简单随机样本，其特定值为$X_1=x_1$，$X_2=x_2$，$\cdots$，$X_n=x_n$.

数学定义的上述两个条件的另一种表述是:(1)n 次试验相互独立,(2)在各次试验中,X 的概率分布保持不变.

上册3.18节比较直观地给出了对总体的**简单随机抽样的非数学定义**:

简单随机抽样是一种抽样方法,对总体的每一次选取,总体中的所有元素(抽样单位)都有相同的被选概率.

在一般的统计学著作中还可以找到两个类似的非数学定义:

来自总体的一个 n 个元素样本的是简单随机样本,如果总体的所有可能的 n 个元素的样本都有相同的被选概率.

来自总体的一个 n 个元素的样本是简单随机样本,如果总体的所有元素都有相等且独立的被选概率.

非数学定义意味着对同一个基本的数学定义给出解释.在本章随后几节的讨论中,如果这些非数学定义的指定条件能够满足,则在一定限制下可以接受抽样是简单随机抽样.

例 13.1　字母表的前9个字母(A至I)被分别打印在完全相同的卡片上,并将卡片放入一只碗中.如果闭上眼睛从碗中一次一张地抽出3张卡片,则这是来自有限总体的无放回抽样(见上册,3.16节).(a)依据非数学定义,这是简单随机抽样吗?(b)用上述方法,从9个字母可以抽取多少个不同的3字母样本?(c)每个3字母样本的抽取概率是多少?

解 ☞　(a)如果在每次选取中所有字母有相等的被抽概率(即,使用上述的第一个非数学定义形式),这样取出3个字母(卡片)构成的样本是一个简单随机样本.

(b)由于与选取次序无关,每个样本是从 $N=9$ 个不同个体(字母)中选取 $n=3$ 个的可能**组合**.因此,由上册公式(9.20),从9个字母中取3个字母的简单随机样本(组合)有

$$_NC_n=\binom{N}{n}=\frac{N!}{n!(N-n)!}=\frac{9!}{3!(9-3)!}=\frac{9\times8\times7}{3\times2\times1}=84$$

(c)对84个简单随机样本的每一个,取出的字母可以有任意次序.第一个字母(L_1)的概率是

$$P(L_1)=\frac{3}{9}$$

给定 L_1 条件下第二个字母(L_2)的概率是

$$P(L_2\mid L_1)=\frac{2}{8}$$

最后,给定 L_1 和 L_2 条件下第三个字母(L_3)的概率是

$$P(L_3\mid L_2\cap L_1)=\frac{1}{7}$$

用上册公式(9.5)计算得到样本($L_1\cap L_2\cap L_3$)的概率:

$$P(L_1\cap L_2\cap L_3)=P(L_1)P(L_2\mid L_1)P(L_3\mid L_2\cap L_1)=\frac{3}{9}\times\frac{2}{8}\times\frac{1}{7}=0.011905$$

例 13.2　在例13.1,如果不是每次取出一个字母,而是将84个可能的样本分别打印在完全相同的卡片上,然后随机选取一张卡片来获得3字母的样本,则:(a)依据非数学定义,这是一个简单随机样本吗?(b)84个可能样本的每一个的被取概率是多少?

解 ☞　(a)依据上述的简单随机样本的非数学定义的第二个形式,由于所有可能的卡片都有相等的被取概率,故这是一个简单随机样本.

(b)从84个可能的3字母样本取出其一的概率是

$$\frac{1}{\binom{N}{n}}=\frac{1}{\binom{9}{3}}=\frac{1}{84}=0.011905$$

与每次一个地取出3个字母得到的概率相同.这表明从有限总体无放回抽样的一般规则:简单随机样本既可以每次抽取一个元素得到,如果总体的所有元素有相等的被选概率;也可以作为一个完整样本一次抽取得到,

如果所有样本有相等的被选概率.

13.4　抽样方法的假定

用推断方法做出有效的概率决策,必须满足一些假定(见上册,3.13节).基于所有推断方法的抽样假定之一是**独立性假定**:从总体中抽取一个元素不影响其它任何元素的抽取概率.推断性统计学的另一个重要假定是**随机抽样假定**:取自总体的是一个简单随机样本.这两个假定并不一样.尽管所有的推断方法都要求样本观测值相互独立,但是一些推断性方法只对于简单随机抽样的概率抽样有效(见上册,3.17节).

当得到一个简单随机样本时,两个假定自然满足,理论上的简单随机样本(见13.3节)要求满足独立性假定,而且,独立性条件至少可以用某一个非数学定义(见13.3节给出的第三个形式)来描述.然而,事实上,如果简单随机样本或者来自无限总体或者来自有放回的有限总体,观测值的独立性假定能够满足;如果简单随机样本来自无放回的有限总体(见例13.1和13.2),独立性假定可能不满足.

对无限大总体的简单随机抽样是理想的抽样情形,因为它满足所有的抽样假定.此时,我们可以认为:(1) 总体分布在抽样过程中保持不变,无论抽样是有放回还是无放回,(2) 从这个不变总体中每次抽样时,总体的每一元素有相同的被取概率,(3) 总体中元素的所有可能组合有相同的被取概率,(4) 总体中所有元素有相同且相互独立的被取概率.

从有限总体有放回的简单随机抽样可看做是从无限总体简单随机抽样.因为在各次选取中元素被放回,总体分布保持不变,从而可以认为满足上述所有来自无限总体的抽样假定.

从有限总体无放回的简单随机抽样违反了独立性假定,因为取走一个元素改变了所有剩余元素的被取概率.然而,对于这样的违例,推断性方法是稳健的(见上册,习题3.14):如果有限总体的容量 N 相对于样本容量 n 较大,则这种违例不构成问题.一些统计学著作指出,除非 n 超过 N 的5% ($n>0.05N$),否则无放回抽样不会构成问题.

在本章及以后几章中,**术语"随机样本"或"样本"即指简单随机样本.抽样是数学定义还是非数学定义,样本来自有限总体还是无限总体,抽样是有放回还是无放回,从上下文即可明了.**

13.5　随机变量 $\bar{X}$

总体以及从总体抽取的样本都有算术平均值.总体均值是由上册公式(3.2)定义的一个参数:

$$\mu=\frac{\sum_{i=1}^{N}x_i}{N}$$

样本均值是由上册公式(3.1)定义的一个统计量:

$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}$$

例13.3　下面是关于数字($N=5$):0,1,2,3,4的5个元素组成的总体.(a) 用上册附录表A.1和例3.5的方法,采取有放回抽样,从该总体取出5个包含2个数字($n=2$)的样本.(b)计算总体的算术平均值和取自总体的5个样本的每一个的算术平均值.(c)从5个样本均值计算总的均值.

解　(a) 首先在上册表A.1中确定一个起始位置:第17行和第37列的交叉处是数51.然后由这个起始位置沿第37列的左边向下,从总体中有放回(即接受重复数字)抽取5个不同数对,如下:(2,2),(4,2),(0,1),(3,2),(2,1).

(b) 由上册公式(3.2),总体的算术平均值是

$$\mu=\frac{0+1+2+3+4}{5}=2.0$$

由上册公式(3.1),样本的算术平均值是

$$\bar{x}_1=\frac{2+2}{2}=2.0, \bar{x}_2=\frac{4+2}{2}=3.0$$

$$\bar{x}_3=\frac{0+1}{2}=0.5, \bar{x}_4=\frac{3+2}{2}=2.5, \bar{x}_5=\frac{2+1}{2}=1.5$$

(c) 上册 6.10 节指出,如果从总体抽样若干次,并对每个样本计算其算术平均值,则这些样本对总体均值的最好估计是用上册公式(6.19)计算出的总平均值:

$$\text{总平均值}=\frac{\sum_{i=1}^{5}\bar{x}_i n_i}{\sum_{i=1}^{5}n_i}$$

$$=\frac{(2\times 2.0)+(2\times 3.0)+(2\times 0.5)+(2\times 2.5)+(2\times 1.5)}{2+2+2+2+2}=1.9$$

这个例子表明,虽然总体只有一个算术平均值 μ,且为常数,但来自总体的容量为 n 的一列样本有各自的算术平均值 $\bar{x}$,并且这些平均值对不同的样本是变化的.如果把每个样本均值看做是一次统计试验随机确定的数值结果,则 $\bar{X}$ 是一个随机变量(见上册,10.1 节).同其它的随机变量一样,必须区分随机变量 $\bar{X}$ 和它的实数值 $\bar{x}$.因此,对例 13.3 的抽样问题,$\bar{X}$ 可取 25 个值(见例 13.4),其中有:$\bar{x}_1=2.0$, $\bar{x}_2=3.0$, $\bar{x}_3=0.5$, $\bar{x}_4=2.5$, $\bar{x}_5=1.5$.

13.6 均值的理论抽样分布和经验抽样分布

如果总体的随机样本是由离散随机变量 X 产生,则 $\bar{X}$(见 13.5 节)也是离散随机变量.同其它变量一样,$\bar{X}$ 有一个由概率累积函数[$f(\bar{x})$]定义的离散型概率分布,$f(\bar{x})$的定义域包含 $\bar{X}$ 的所有可能取值($X=\bar{x}$),其值域是这些取值的相应概率[$P(\bar{X}=\bar{x})$](见上册,10.3 节).

如果总体的随机样本是由连续随机变量 X 产生,则 $\bar{X}$(见 13.5 节)也是连续随机变量.同其它变量一样,$\bar{X}$ 有一个由概率密度函数[$f(\bar{x})$]定义的连续型概率分布,$f(\bar{x})$的定义域由 $\bar{X}$ 的无限或不可数个取值($\bar{X}=\bar{x}$)组成,其值域是这些取值的概率密度函数.

均值的理论抽样分布是一个概率分布(离散型或连续型),由与 $\bar{X}$ 的所有取值对应的$f(\bar{x})$值(概率或密度)组成.被称为均值的抽样分布是因为,它代表着从总体抽取固定容量 n 的重复样本的均值.而被称为理论抽样分布是因为,它包含所有的可能均值,并且这种包含通常要求从一个理论模型进行数学推导.

相比较,**均值的经验抽样分布**确实来自总体的固定容量 n 的重复随机样本,并对每一样本计算算术平均值 $\bar{X}=\bar{x}$,然后构造这些平均值的相对频数分布.回顾 10.5 和 10.6 节,概率分布是关于总体相对频数分布的数学模型,而且,随着样本容量的增加,样本相对频数分布越来越接近概率分布.在抽样理论中,概率分布是关于总体经验抽样分布的数学模型,而且随着容量为 n 的样本个数向着不可能达到的目标——来自总体的"所有可能样本"——增加,均值的经验抽样分布越来越接近均值的理论抽样分布.

理论抽样分布适用于上册第 6 章和第 7 章讨论的大部分描述性统计量,而且可以用合适的经验分布来近似.例如,来自容量为 N 的总体的一个容量为n 的样本,样本方差 s^2(见上册,7.7 节)实际上是一个可取特定值 s^2($S^2=s^2$)的随机变量 S^2;而且,方差存在理论抽样分布,并能用方差的经验抽样分布近似.本章主要讨论均值的抽样分布,但是也将考虑**样本和**(见 13.19 节),**成功次数**(见 13.22 节)和**比例**(见 13.23 节)**的抽样分布**.其它重要的抽样分布,将在下几章介绍,包括 ***t* 分布**(第 14 章),**χ^2 分布**(第 15 章),***F* 分布**(第 17 章)和**两均值差的抽样分布**(第 17 章).

典型的统计分析往往要涉及到总体的未知参数和一个用于对参数值进行推断的样本[即,估计问题或假设检验问题(见上册,3.6 节)].用于估计参数值的样本统计量的值容易计算,但是这个估计有多好——它与参数真值的接近程度如何?样本是一个"典型的"随机样本还是一个极端形式?样本统计量的值在重复抽样过程中变化幅度是否大?在推断性统计学中,这些以及其它的这类问题将由理论抽样分布得到回答.实际上,理论抽样分布是各种统计推断的基

础.

例 13.4　对例 13.3 的数字:0,1,2,3,4 的总体,用有放回抽样得到一个容量 $n=2$ 的随机样本,确定其均值的离散型理论抽样分布.按以下步骤来确定该分布:(a)列出从总体抽取的所有可能的样本及每个样本的算术平均值.(b)用频数和相对频数分布概述这些算术平均值.(c)计算每一均值的被取概率,并且分别用概率表和概率直方图(见上册,10.3 节)表示得到的抽样分布.

解 ☞　(a) 由计数规则:乘法原理(见上册,9.12 节)知,如果有放回抽样,来自该总体的容量 $n=2$ 的可能随机样本有($n_1\times n_2=5\times5=25$)个.这些样本及其均值[(样本)$\bar{x}$]是:

[(0, 0)0.0]　[(1, 0)0.5]　[(2, 0)1.0]　[(3, 0)1.5]　[(4, 0)2.0]
[(0, 1)0.5]　[(1, 1)1.0]　[(2, 1)1.5]　[(3, 1)2.0]　[(4, 1)2.5]
[(0, 2)1.0]　[(1, 2)1.5]　[(2, 2)2.0]　[(3, 2)2.5]　[(4, 2)3.0]
[(0, 3)1.5]　[(1, 3)2.0]　[(2, 3)2.5]　[(3, 3)3.0]　[(4, 3)3.5]
[(0, 4)2.0]　[(1, 4)2.5]　[(2, 4)3.0]　[(3, 4)3.5]　[(4, 4)4.0]

(b) 频数和相对频数分布由表 13.1 给出,这里 N_r 表示可能的 $\bar{x}$ 值的个数.

(c) 在随机抽样中,由于所有可能样本都有相等的被取概率(见 13.3 节),25 个可能样本的每一个的被取概率均为:1/25=0.04.这些样本是互斥事件,由概率的集合论解释(见上册,8.6 节)的性质(4)知,选取 k 个有相等均值的样本($S_1,S_2,\cdots,S_k$)的任一个的概率是

$$P(S_1\cup S_2\cup\cdots\cup S_k)=P(S_1)+P(S_2)+\cdots+P(S_k)$$

因此,为得到给定均值的概率,只须将所有样本的概率相加即可.于是,

$$P(\bar{X}=0.0)=0.04$$

$$P(\bar{X}=0.5)=0.04+0.04=0.08$$

$$P(\bar{X}=1.0)=0.04+0.04+0.04=0.12$$

等等,表 13.2 给出了均值的理论抽样分布,图 13-1 是直方图.注意:相对频数分布(见表 13.1)和理论抽样分布(见表 13.2 和图 13-1)相同.

表 13.1

样本均值 $\bar{x}_i$	频数 f_i	相对频数 f_i/N_r
0.0	1	0.04
0.5	2	0.08
1.0	3	0.12
1.5	4	0.16
2.0	5	0.20
2.5	4	0.16
3.0	3	0.12
3.5	2	0.08
4.0	1	0.04
$\sum$	25	1.00

表 13.2

样本均值 $\bar{x}$	概率 $f(\bar{x})$
0.0	0.04
0.5	0.08
1.0	0.12
1.5	0.16
2.0	0.20
2.5	0.16
3.0	0.12
3.5	0.08
4.0	0.04
$\sum$	1.00

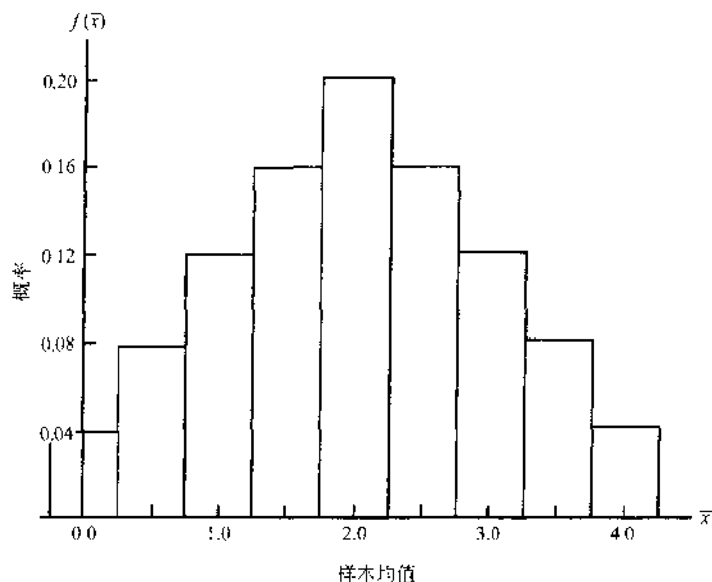


图 13.1

例 13.5　对例 13.3 的关于数字:0,1,2,3,4 的总体,无放回抽取一个容量 $n=2$ 的随机样本,计算样本均值的离散型理论抽样分布.按以下步骤来确定该分布:(a)列出从总体抽取的所有可能的样本及每个样本的算术平均值.(b)用频数和相对频数分布概述这些算术平均值.(c)计算每一个均值的被取概率,并且分别用概率表和概率直方图表示得到的抽样分布.

解 ☞　(a)本题是从有限总体无放回随机抽样,且不考虑取值次序.例如,认为样本(0,4)与(4,0)等同.因此,来自容量为 N 的有限总体的容量为 n 的可能随机样本的总个数,由组合规则[上册,公式(9.20)]计算,

$$_NC_n=\binom{N}{n}=\frac{N!}{n!(N-n)!}$$

代入 $N=5$ 和 $n=2$,可能样本的总个数是

$$_5C_2=\binom{5}{2}=\frac{5!}{2!(5-2)!}=\frac{5\times 4}{2\times 1}=10$$

这些样本及其均值[(样本)$\bar{x}$]是

[(0,1)0.5][(1,3)2.0]

[(0,2)1.0][(1,4)2.5]

[(0,3)1.5][(2,3)2.5]

[(0,4)2.0][(2,4)3.0]

[(1,2)1.5][(3,4)3.5]

(b) 频数和相对频数分布由表 13.3 给出,这里 $N_{\bar{x}}$ 同样表示可能的 $\bar{x}$ 值的个数.

表 13.3

样本均值 $\bar{x}_i$	频数 f_i	相对频数 $f_i/N_{\bar{x}}$
0.5	1	0.1
1.0	1	0.1
1.5	2	0.2
2.0	2	0.2
2.5	2	0.2
3.0	1	0.1
3.5	1	0.1
$\sum$	10	1.0

(c) 概率计算的步骤与例 13.4(c)相同.每一可能样本的概率值已知,从而均值的概率是 1/10=0.1,于是

$$P(\bar{X}=0.5)=0.1$$

$$P(\bar{X}=1.0)=0.1$$

$$P(\bar{X}=1.5)=0.1+0.1=0.2$$

等等.表 13.4 给出了均值的理论抽样分布,图 13-2 是直方图.再次注意:同例 13.4 一样,相对频数分布(见表 13.3)和理论抽样分布(见表 13.4 和图 13-2)相同.

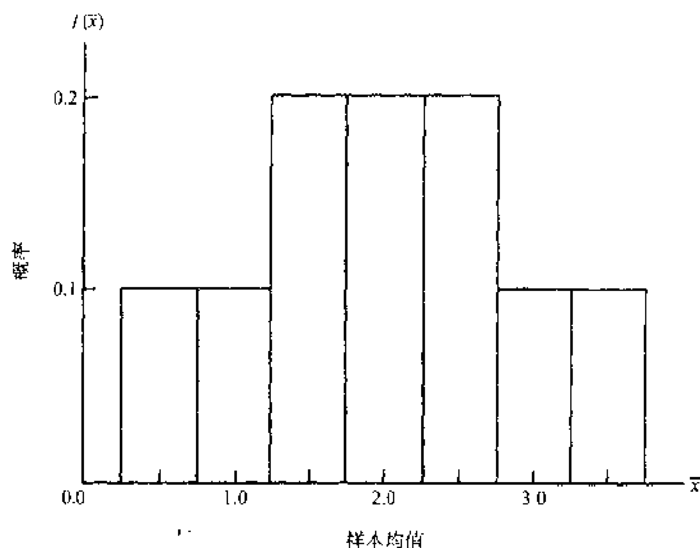


图 13-2

表 13.4

样本均值 $\bar{x}$	概率 $f(x)$
0.5	0.1
1.0	0.1
1.5	0.2
2.0	0.2
2.5	0.2
3.0	0.1
3.5	0.1
$\sum$	1.0

13.7 均值的抽样分布的均值

均值的理论抽样分布是一个概率函数(离散型或连续型),包含对 $\bar{X}$ 的所有可能取值(见 13.6 节)对应的 $f(\bar{x})$值(概率或密度).对均值的任何理论抽样分布,无论离散型或连续型,在数学上可以证明,理论抽样分布的均值 $\mu_{\bar{x}}$[或期望 $E(\bar{X})$]总是等于抽样总体的均值 μ.即,

$$E(\bar{X})=\mu_{\bar{x}}=\mu \tag{13.1}$$

例 13.6 利用上册 10.9 节的 $E(X)$定义,说明对于(a)表 13.2,(b)表 13.4 的理论抽样分布,有 $\mu_{\bar{x}}=\mu$.

解 (a) 上册 10.9 节指出,设 X 是离散随机变量,分别以概率 $f(x_1)$, $f(x_2)$,$\cdots$,$f(x_k)$取值 $x_1,x_2,\cdots,x_k$,则 X 的期望,记为 $E(X)$,是[上册,公式(10.10)]

$$E(X) = \mu = \sum_{i=1}^{k} x_i f(x_i) = \sum_x x f(x)$$

为求 $E(\bar{X})$,将公式修改为

$$E(\bar{X}) = \mu_{\bar{x}} = \sum_{i=1}^{k} \bar{x}_i f(\bar{x}_i) = \sum_{\bar{x}} \bar{x} f(\bar{x}) \tag{13.2}$$

对表 13.2 的分布,由上式,

$$\begin{aligned} E(\bar{X}) = \mu_{\bar{x}} &= (0.0 \times 0.04) + (0.5 \times 0.08) + (1.0 \times 0.12) \\ &\quad + (1.5 \times 0.16) + (2.0 \times 0.20) + (2.5 \times 0.16) \\ &\quad + (3.0 \times 0.12) + (3.5 \times 0.08) + (4.0 \times 0.04) \\ &= 2.0 \end{aligned}$$

这个值与总体均值 $\mu = 2.0$[见例 13.3(b)]相等.

$$\begin{aligned} \text{(b)} E(\bar{X}) = \mu_{\bar{x}} &= (0.5 \times 0.1) + (1.0 \times 0.1) + (1.5 \times 0.2) + (2.0 \times 0.2) \\ &\quad + (2.5 \times 0.2) + (3.0 \times 0.1) + (3.5 \times 0.1) \\ &= 2.0 \end{aligned}$$

13.8 估计量的准确度

用于估计总体参数的样本统计量称为**估计量**.用于表示参数及其估计量的典型符号是:用 θ(小写希腊字母 theta)表示参数,$\hat{\theta}$ 表示其估计量.用 $\hat{\theta}^*$ 表示 $\hat{\theta}$ 的特定估计或数值.

上册 2.14 节指出,在统计学中,某次测量的准确度是指它与真值的接近程度,并且准确度基本上由是否存在系统误差或偏差(见上册,2.13 节)决定.类似地,**估计量 $\hat{\theta}$ 的准确度**是由估计量的理论抽样分布的均值[期望 $E(\hat{\theta})$]与待估的总体参数 θ 的接近程度决定.因此,估计量的准确度的度量是 $|E(\hat{\theta}) - \theta|$,即 $E(\hat{\theta})$ 和 θ 之间距离的绝对值.

$E(\hat{\theta}) - \theta$ 称为**估计量 $\hat{\theta}$ 的偏差**,同测量偏差一样,这个偏差被认为是系统误差的结果.当 $E(\hat{\theta}) = \theta$ 或 $E(\hat{\theta}) - \theta = 0$ 时,估计量称为是**无偏的**或具有**无偏性**.即如果 $\hat{\theta}$ 是无偏的,则它是一个完全准确的估计量.

从例 13.6 知 $E(\bar{X}) = \mu$,所以估计量 $\hat{\theta} = \bar{X}$ 是参数 $\theta = \mu$ 的一个完全准确(无偏)的估计量.

13.9 均值的抽样分布的方差:无限总体或有放回抽样

当抽样是针对无限总体或有放回的有限总体时,在数学上可以证明,对容量为 n 的所有可能的随机样本,样本均值的抽样分布的方差是

$$E[(\bar{X} - \mu_{\bar{x}})^2] = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \tag{13.3}$$

这里,抽样总体的容量为 N,且有均值 μ 和方差 σ^2.

例 13.7 用上册 10.11 节 $E[(X - \mu)^2]$ 的定义,说明表 13.2 给出的均值的理论抽样分布有 $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$.

解 上册 10.11 节指出,设 X 是离散随机变量,分别以概率 $f(x_1)$, $f(x_2)$, $\cdots$, $f(x_k)$ 取值 x_1, x_2, $\cdots$, x_k,则 X 的方差是

$$E[(X - \mu)^2] = \sigma^2 = \sum_x (x - \mu)^2 f(x)$$

为求 $E[(\bar{X} - \mu_{\bar{x}})^2]$,将公式修改为

$$E[(\bar{X} - \mu_{\bar{x}})^2] = \sigma_{\bar{x}}^2 = \sum_{\bar{x}} (\bar{x} - \mu_{\bar{x}})^2 f(\bar{x}) \tag{13.4}$$

对表 13.2 中分布，$\mu_{\bar{x}}=2.0$［见例 13.6(a)］，由上式

$$E[(\bar{X}-\mu_{\bar{x}})^2]=[(0.0-2.0)^2 0.04]+[(0.5-2.0)^2 0.08]+[(1.0-2.0)^2 0.12]$$
$$+[(1.5-2.0)^2 0.16]+[(2.0-2.0)^2 0.20]+[(2.5-2.0)^2 0.16]$$
$$+[(3.0-2.0)^2 0.12]+[(3.5-2.0)^2 0.08]+[(4.0-2.0)^2 0.04]$$
$$=1.00$$

为说明这个值等于$\frac{\sigma^2}{n}$，对总体：0，1，2，3，4 计算 σ^2，将 $\mu=2.0$［见例 13.6(a)］代入上册公式(7.12)：

$$\sigma^2=\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$$
$$=\frac{(0-2.0)^2+(1-2.0)^2+(2-2.0)^2+(3-2.0)^2+(4-2.0)^2}{5}=2.00$$

因此，当 $n=2$ 时，

$$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}=\frac{2.00}{2}=1.00$$

13.10 均值的抽样分布的方差：无放回抽样的有限总体

如果对有限总体无放回抽样，在数学上可以证明，对容量为 n 的所有可能的随机样本，样本均值的抽样分布的方差是

$$E[(\bar{X}-\mu_{\bar{x}})^2]=\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{N-1} \tag{13.5}$$

这里，抽样总体的容量为 N，且有均值 μ 和方差 σ^2．在上式中，因子$\frac{N-n}{N-1}$称为**有限总体的修正因子平方**．

例 13.8 用公式(13.4)说明：表 13.4 中均值的理论抽样分布有 $\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{N-1}$．

解 ☞ 由公式(13.4)

$$E[(\bar{X}-\mu_{\bar{x}})^2]=\sigma_{\bar{x}}^2=\sum_{\bar{x}}(\bar{x}-\mu_{\bar{x}})^2 f(\bar{x})$$

因此，对表 13.4 中分布，$\mu_{\bar{x}}=2.0$［见例 13.6(a)］，有

$$\sigma_{\bar{x}}^2=[(0.5-2.0)^2 0.1]+[(1.0-2.0)^2 0.1]+[(1.5-2.0)^2 0.2]+[(2.0-2.0)^2 0.2]$$
$$+[(2.5-2.0)^2 0.2]+[(3.0-2.0)^2 0.1]+[(3.5-2.0)^2 0.1]$$
$$=0.75$$

为说明这个值等于$\frac{\sigma^2}{n}\times\frac{N-n}{N-1}$，代入例 13.7 的 $n=2$、$N=5$ 和总体方差 $\sigma^2=2.00$

$$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{N-1}=\frac{2.0}{2}\times\frac{5-2}{5-1}=0.75$$

如果抽样对无放回的有限总体，计算均值的抽样分布的方差可以忽略有限总体的修正因子平方，只要总体容量 N 相对于样本容量 n 较大．这是因为，当 n 保持不变时，随着 N 增加，$\frac{N-n}{N-1}$趋近于它的极限 1.0．因此，许多统计学著作给出一个通用规则：当 n 小于或等于 N 的 5%（$n\leqslant 0.05N$）时，不必使用修正因子平方

13.11 均值的标准误

如果从总体抽取一个随机样本，并且从该样本计算出估计量 $\hat{\theta}$ 的具体值 $\hat{\theta}^*$（见 13.8 节），则 $\hat{\theta}^*$ 和待估参数 θ 的差 $\hat{\theta}^*-\theta$ 称为**抽样误差**．如果估计量是无偏的，抽样误差也称为

随机误差,此时,抽样误差来源于不同样本估计值的随机影响.也就是说,对这样一个随机样本,如果用计算出的样本均值($\bar{X}=\bar{x}$)来估计总体参数 μ,则这个估计值和 μ 之间的差($\bar{x}-\mu$)称为抽样误差或随机误差($\bar{X}$ 是 μ 的无偏估计).

回顾上册 7.9 节,总体方差的正平方根($\sqrt{\sigma^2}=\sigma$)称为标准差,因为它是关于总体均值的任意度量的标准或典型偏差($x_i-\mu$,见上册,6.4 节)的一个度量.类似地,均值的抽样分布的方差的正平方根

$$\sqrt{E[(\bar{X}-\mu_{\bar{x}})^2]}=\sigma_{\bar{x}}=\sqrt{\sigma_{\bar{x}}^2} \tag{13.6}$$

称为**均值的标准误**,因为它是关于总体均值的标准或典型的抽样误差的一个度量.

如果是对无限总体或有放回的有限总体抽样,有[公式(13.3)]

$$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}$$

因此,在这样的抽样条件下,均值的标准误是

$$\sigma_{\bar{x}}=\sqrt{\sigma_{\bar{x}}^2}=\sqrt{\frac{\sigma^2}{n}}=\frac{\sqrt{\sigma^2}}{\sqrt{n}}=\frac{\sigma}{\sqrt{n}} \tag{13.7}$$

例 13.9 对表 13.2 的均值的理论抽样分布,分别利用抽样分布的方差 $\sigma_{\bar{x}}^2$ 的平方根和公式(13.7)计算均值的标准误.

解 ☞ 由例 13.7 知,$\sigma_{\bar{x}}^2=1.00$.因此,均值的标准误是

$$\sigma_{\bar{x}}=\sqrt{\sigma_{\bar{x}}^2}=\sqrt{1.00}=1.00$$

将 $\sigma=\sqrt{2.00}$(见例 13.7)代入公式(13.7)计算 $\sigma_{\bar{x}}$,

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{\sqrt{2.00}}{\sqrt{2}}=1.00$$

对于有限总体的无放回抽样,有[公式(13.5)]

$$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{N-1}$$

从而,在这些抽样条件下,均值的标准误是

$$\sigma_{\bar{x}}=\sqrt{\sigma_{\bar{x}}^2}=\sqrt{\frac{\sigma^2}{n}\times\frac{N-n}{N-1}}=\sqrt{\frac{\sigma^2}{n}}\sqrt{\frac{N-n}{N-1}}=\frac{\sigma}{\sqrt{n}}\times\sqrt{\frac{N-n}{N-1}} \tag{13.8}$$

这里 $\sqrt{\frac{N-n}{N-1}}$ **是有限总体的修正因子**.

例 13.10 对表 13.4 的均值的理论抽样分布,分别利用抽样分布的方差 $\sigma_{\bar{x}}^2$ 的平方根和公式(13.8)计算均值的标准误.

解 ☞ 由例 13.8 知,$\sigma_{\bar{x}}^2=0.75$.因此,均值的标准误是

$$\sigma_{\bar{x}}=\sqrt{\sigma_{\bar{x}}^2}=\sqrt{0.75}=0.8660$$

将 $\sigma=\sqrt{2.00}$(见例 13.7)代入公式(13.8)计算 $\sigma_{\bar{x}}$,

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}\times\sqrt{\frac{N-n}{N-1}}=\frac{\sqrt{2.00}}{\sqrt{2}}\times\sqrt{\frac{5-2}{5-1}}=1.00\sqrt{0.75}=0.8660$$

所有的描述性统计量都是作为总体参数的估计量提出的.对大部分这样的统计量,存在理论抽样分布(对所有相同容量的随机样本),因为这些统计量是对应参数的无偏估计(见 13.8 节),所以它们的理论抽样分布的均值等于它们所估计的参数.统计量的某个计算值与待估参数的差是抽样误差.因此,对任何给定的统计量,理论抽样分布的标准差称为该统计量的标准

误——统计量关于它的参数随不同样本的变化程度.在本章和下一章,当介绍新的理论抽样分布时,会给出统计量的标准误以及相应计算公式.

13.12 估计量的精密度

上册 2.14 节介绍了测量的两个基本统计性质:准确度(测量值与真实值的接近程度)和精密度(对同一子样重复测量的接近程度).类似地,准确度[用 $|E(\hat{\theta})-\theta|$(见 13.8 节)度量]和精密度也是估计量 $\hat{\theta}$ 的两个基本性质.

如果容量为 n 的随机样本重复的取自容量为 N 的总体,并且为估计总体参数 μ,从每个样本计算出估计量 $\hat{\theta}$ 的各个值 $\hat{\theta}'$,则**估计量的精密度**(简称**精度**)由重复估计的变异性或离散程度决定.估计值的变化越小(聚集越紧密),估计量越精密.

从这个定义可以看出,对于估计量 $\hat{\theta}$(见 13.11 节),抽样分布的标准误是估计量 $\hat{\theta}$ 的精密度的一个度量,标准误越小,估计量越精密,因此,$\bar{X}$ 作为 μ 的估计量,均值的标准误 $\sigma_{\bar{x}}$ 就是 $\bar{X}$ 的精密度的度量;$\sigma_{\bar{x}}$ 越小,$\bar{X}$ 估计 μ 越精密.如果 $\hat{\theta}$ 是一个无偏估计(见 13.8 节),例如 $\bar{X}$ 是 μ 的无偏估计,则说标准误度量**随机误差**——由于随机抽样引起的一列估计值的变异.如果存在两个无偏估计 $\hat{\theta}_1$ 和 $\hat{\theta}_2$,$\hat{\theta}_1$ 比 $\hat{\theta}_2$ 有更小的标准误,则称 $\hat{\theta}_1$ 更精密.

作为 μ 的一个估计,$\bar{X}$ 的精密度由均值的标准误度量.从 $\sigma_{\bar{x}}$ 的两个不同公式[公式(13.7)和(13.8)]均可看出:如果 n 增加,而 σ 保持不变,则 $\sigma_{\bar{x}}$ 减小从而 $\bar{X}$ 作为 μ 的估计将更精密.反之亦然.如果 n 保持不变,而 σ 增加,则 $\sigma_{\bar{x}}$ 增加,从而 $\bar{X}$ 精度将减小.

例 13.11 从无限总体抽取一个容量为 n 的随机样本,总体标准差 $\sigma=2.0$,试说明:如果样本容量从 $n=4$ 增加到 $n=16$,$\bar{X}$ 作为 μ 的估计将更加精密.

解 作为 μ 的估计,$\bar{X}$ 的精密度由均值的标准误 $\sigma_{\bar{x}}$ 度量.从无限总体抽样,$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}$.因此,代入 $\sigma=2.0$ 和 $n=4$.

$$\sigma_{\bar{x}}=\frac{2.0}{\sqrt{4}}=\frac{2.0}{2}=1.0$$

n 从 4 增加到 16,

$$\sigma_{\bar{x}}=\frac{2.0}{\sqrt{16}}=\frac{2.0}{4}=0.5$$

于是,如果样本容量增加 4 倍,而 σ 不变,$\sigma_{\bar{x}}$ 将减少 50%.

13.13 用均值的离散型抽样分布计算概率

上册 10.7 节指出,利用离散随机变量 X 的累积分布函数可以计算 X 取值小于等于 x 的概率.类似地,离散随机变量 $\bar{X}$ 的累积分布函数也给出 $\bar{X}$ 取值小于等于 $\bar{x}$ 的概率.用 $F(\bar{x})$ 表示这个函数,对任意实数($-\infty<\bar{x}<\infty$)

$$F(\bar{x})=P(\bar{X}\leqslant\bar{x}) \tag{13.9}$$

对任意实数 a,由下式计算 $P(\bar{X}\leqslant a)$

$$F(a)=\sum_{\bar{x}\leqslant a}f(\bar{x}) \tag{13.10}$$

这里,符号 $\sum_{\bar{x}\leqslant a}f(\bar{x})$ 表示对任意小于等于 a 的 $\bar{x}$ 值对 $f(\bar{x})$ 求和.

例 13.12 对例 13.3 的数字:0,1,2,3,4 的总体以及表 13.2 给出的均值的离散型抽样分布,从该总体有放回抽取容量为($n=2$)的随机样本,计算样本均值($\bar{X}=\bar{x}$)至多为 2 的概率.

解 由公式(13.10)知

$$F(2.0)=\sum_{\bar{x}\leqslant 2.0}f(\bar{x})$$

查表13.2找到所需的 $f(\bar{x})$ 值,代入得

$$\begin{aligned}F(2.0) &= f(0.0) + f(0.5) + f(1.0) + f(1.5) + f(2.0)\\ &= 0.04 + 0.08 + 0.12 + 0.16 + 0.20\\ &= 0.60\end{aligned}$$

13.14 用均值的正态抽样分布计算概率

在理论上,称一个总体无限大且服从正态分布,如果它的连续随机变量 X 有正态概率分布(见12.2节).在实际应用中,称一个任意容量的总体服从正态分布,如果它的经验频数(或相对频数,也称频率)分布能用一个正态曲线合理的拟合.由理论定义(从而对应用定义也成立),在数学上可以证明:如果容量为 n 的所有可能的随机样本均来自一个无限大的正态分布总体(均值 μ,方差 σ^2,标准差 σ),并且对每一样本计算连续随机变量 X 的取值 $\bar{x}$,则 $\bar{X}$ 有一个服从正态分布的抽样分布,其均值 $\mu_{\bar{x}} = \mu$,方差 $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$,标准差 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.这个结果对所有容量为 n 的样本严格成立.

这样的一个服从正态分布的连续随机变量 $\bar{X}$ 能用 Z 变换[公式(12.3)]进行标准化

$$Z = \frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \tag{13.11}$$

它将 $\bar{X}$ 的正态分布变换为由公式(12.5)定义的标准正态分布($\mu_z = 0$, $\sigma_z^2 = \sigma_z = 1$).因为从无限大总体抽样,$\bar{X}$ 的分布(见13.7和13.9节)有 $\mu_{\bar{x}} = \mu$ 和 $\sigma_{\bar{x}} = \frac{\sigma}{n}$,所以,$Z$ 变换的以上形式在一些教科书也写成

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{x}}} \tag{13.12}$$

和

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \tag{13.13}$$

例 13.13 从正态分布总体($\mu = 20.0$, $\sigma = 1.0$)抽取一个随机样本($n = 100$).(a)连续随机变量 X 的分布的特征量(即形状和参数)是什么?(b)$\bar{X}$ 在区间 $20.0 \leqslant \bar{x} \leqslant 20.2$ 上取值的概率是多少?

解 (a) $\bar{X}$ 服从正态分布,且 $\mu_{\bar{x}} = 20.0$ 和 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.0}{\sqrt{100}} = 0.10$.

(b) 如果用公式(13.12)标准化 X,则得到的连续变量 Z 服从标准正态分布.因此,公式(12.7)可以改写为

$$P(a \leqslant \bar{X} \leqslant b) = P\left(\frac{a - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \leqslant Z \leqslant \frac{b - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) = P(z_a \leqslant Z \leqslant z_b) \tag{13.14}$$

由附录表A.5和公式(13.14),

$$\begin{aligned}P(20.0 \leqslant \bar{X} \leqslant 20.2) &= P\left(\frac{20.0 - 20.0}{0.10} \leqslant Z \leqslant \frac{20.2 - 20.0}{0.10}\right)\\ &= P(0.00 \leqslant Z \leqslant 2.00)\\ &= 0.4772\end{aligned}$$

13.15 中心极限定理:从有限总体有放回抽样

13.6节指出,理论抽样分布实际上是各种统计推断的基础.然而,这些理论工具仅仅适用于分布特征(形状,参数)已知的情形.

在假定抽样总体服从正态分布的条件下,13.14节已经清楚的表述了均值 $\bar{X}$ 的理论抽样

分布的特征.当时指出,在数学上可以证明,如果容量为 n 的所有的可能样本取自正态分布总体,并且从每个样本计算出 $X=\bar{x}$,则得到的均值的连续型理论抽样分布服从正态分布.进一步指出,这个结果对所有容量为 n 的样本严格成立.

但是,在统计学的实际应用中,如果总体分布的形状未知,或者总体具有偏态或多峰,或者总体是有限的离散随机变量,情况又如何呢?值得注意的是,对于应用中遇到的所有这类总体分布,由此导出的理论抽样分布可以认为近似服从正态分布,如果样本容量"足够大"("足够大"的意义见 13.18 节).这是推断性统计学的一个最重要的定理——**中心极限定理**的基础.这个定理,首先由 Laplace 发现(见 12.2 节),对有限总体有放回抽样,定理表述为:在数学上可以证明,如果所有容量为 n 的随机样本,均有放回地取自容量为 N,具有有限参数(μ, σ^2, σ)的有限总体,并对每一样本计算出 $\bar{X}=\bar{x}$,则如果 n 足够大,得到的均值的理论抽样分布近似服从参数为 $\mu_{\bar{x}}=\mu$, $\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}$ 和 $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}$ 的正态分布.

这样,无论被抽样的有限总体的分布特征如何,如果样本容量 n 足够大,则得到的均值的抽样分布近似服从正态分布,并且随着 n 增加,这种近似越来越好.这个定理适用于任何有放回抽样的有限总体.

例 13.14　试验是从 3 个元素($N=3$):3,4,5 的总体中随机选取一个数.随机变量"被选数字"的概率分布是离散型均匀分布(见 12.14 节),如图 13-3 显示.用例 13.4 的方法,给出均值的离散型理论抽样分布:(a)有放回地随机选出 2 个数字,(b)有放回地随机选出 4 个数字.然后(c)对每个分布计算 $\mu_{\bar{x}}$ 和 $\sigma_{\bar{x}}$.

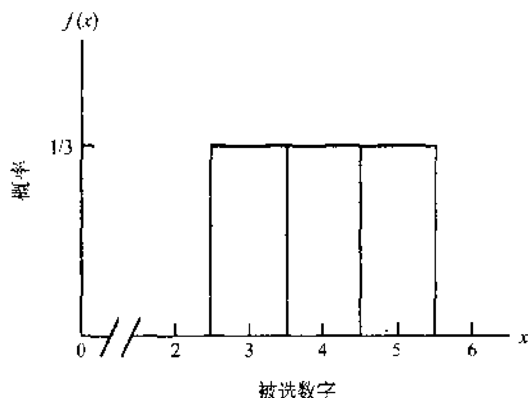


图 13-3

解　(a) 有($3\times3=9$)个容量为 $n=2$ 的可能样本,产生 5 个可能的样本均值.表 13.5 给出均值的理论抽样分布,图 13-4 是直方图.

(b) 有($3\times3\times3\times3=81$)个容量为 $n=2$ 的可能样本,产生 9 个可能的样本均值.表 13.6 给出均值的理论抽样分布,图 13-5 是直方图.

(c) 对上述总体,

$$\mu=\frac{\sum_{i=1}^{N}x_i}{N}=\frac{3+4+5}{3}=4.0$$

表 13.5

样本均值 $\bar{x}$	概率 $f(\bar{x})$
3.0	0.111111
3.5	0.222222
4.0	0.333333
4.5	0.222222
5.0	0.111111
$\sum$	0.999999

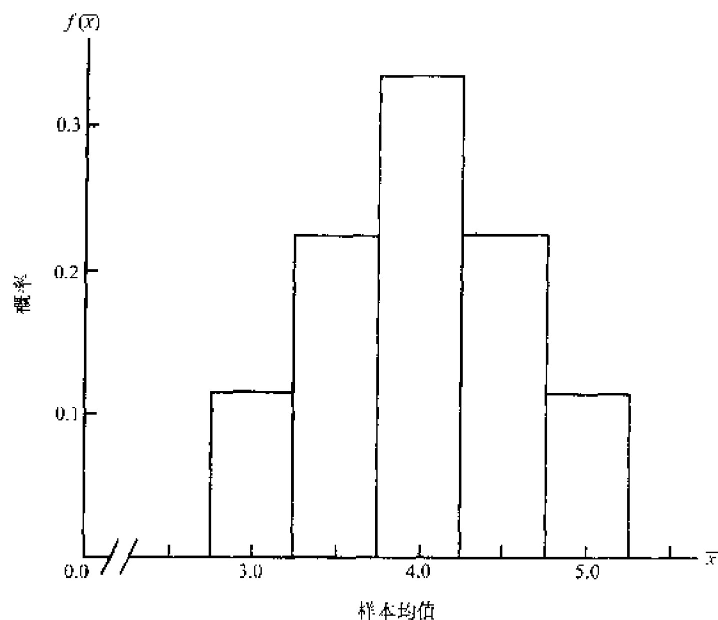


图 13-4

且

$$\sigma=\sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$$

$$=\sqrt{\frac{(3-4.0)^2+(4-4.0)^2+(5-4.0)^2}{3}}$$

$$=0.816497 \text{ 或 } 0.82$$

因此,对(a)中的分布,由公式(13.1)和(13.7),

$$\mu_{\bar{x}}=\mu=4.0$$

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{0.816497}{\sqrt{2}}=\frac{0.816497}{1.414214}=0.577350 \text{ 或 } 0.58$$

表 13.6

样本均值 $\bar{x}$	概率 $f(\bar{x})$
3.00	0.012346
3.25	0.049383
3.50	0.123457
3.75	0.197531
4.00	0.234568
4.25	0.197531
4.50	0.123457
4.75	0.049383
5.00	0.012346
$\sum$	1.000002

对(b)中的分布,

$$\mu_{\bar{x}}=\mu=4.0$$

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{0.816497}{\sqrt{4}}=\frac{0.816497}{2}=0.408249 \text{ 或 } 0.41$$

这个例子说明了中心极限定理的正确性.这里是对有限总体($N=3$)有放回抽样,并且总

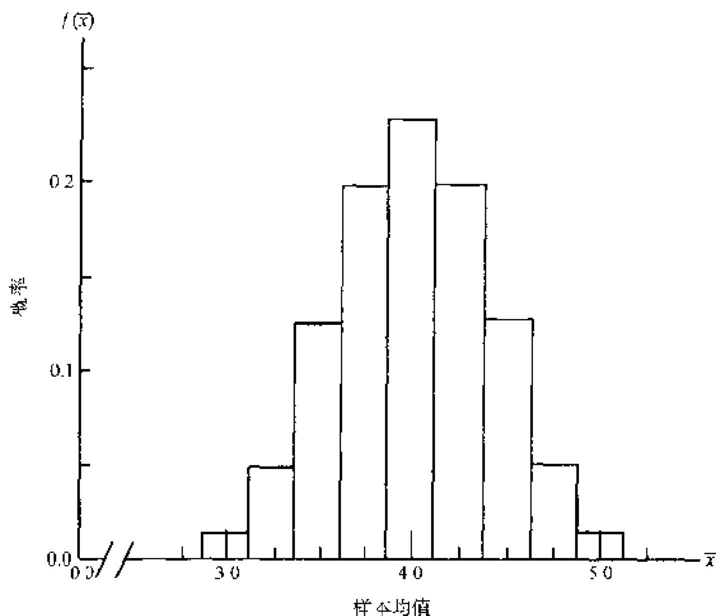


图 13-5

体有离散型均匀概率分布(见图 13-3).即使基于这样一个分布,由于样本容量从 $n=2$ 增加到 $n=4$,均值的抽样分布也变得更接近于正态分布(单峰,对称,钟形;见图 13-4 和 13-5).注意到,随着 n 增加,$\sigma_{\bar{x}}$ 减小(从 0.58 到 0.41);这表明,就像正态分布一样,概率更集中在均值附近.

13.16　中心极限定理:从无限总体抽样

13.15 节描述了中心极限定理如何应用于来自有限总体的有放回抽样的样本.中心极限定理也适用于任何无限大总体的抽样:在数学上可以证明,如果所有容量为 n 的随机样本均取自(有放回或无放回)一个有有限参数(μ,σ^2,σ)的无限总体,并对每一样本计算出 $\bar{X}=\bar{x}$,则如果 n 足够大,得到的均值的理论抽样分布近似服从参数为 $\mu_{\bar{x}}=\mu$,$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}$ 和 $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}$ 的正态分布.

值得注意:即使总体无限大,正态分布拟合均值的抽样分布的好坏程度仍然由样本容量 n 而不是总体容量 N 决定.

13.17　中心极限定理:从有限总体无放回抽样

中心极限定理的第三种形式适用于从有限总体进行无放回抽样.在这种形式中,同 13.10 和 13.11 节的讨论一样,要求 N 对 n 有一定关系.定理表述为:在数学上可以证明,如果所有容量为 n 的随机样本均无放回地取自容量为 N,具有有限参数(μ,σ^2,σ)的有限总体,并对每一样本计算出 $\bar{X}=\bar{x}$,而且 N 至少是 n 的两倍($N\geqslant 2n$),则如果 n 足够大,均值的理论抽样分布近似服从参数为 $\mu_{\bar{x}}=\mu$,$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}\times\frac{N-n}{N-1}$ 和 $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}\times\sqrt{\frac{N-n}{N-1}}$ 的正态分布.

回顾 13.10 节,如果 $n\leqslant 0.05N$,不必使用有限总体的修正因子.

13.18　多大是"足够大?"

在中心极限定理的上述三种形式中,均要求:如果 n"足够大",均值的理论抽样分布近似服从正态分布.在中心极限定理中,要求 n"足够大"是因为:使用中心极限定理所要求的样本

容量 n 是总体分布形状的一个函数，是变化的；对"足够大"不存在绝对的统一规则．总体分布越接近于正态(单峰，对称，钟形)，所必需的样本容量 n 越小．对于这类"正态"总体，一些统计学著作认为，$n=25$ 或 $n=20$ 的样本即足够大．总体越偏，要求的样本越大．对任何类型的总体分布，通常可接受的规则是：如果 $n\geqslant 30$，即认为样本容量足够大，可使用中心极限定理，并能达到合理的准确度．

因为这个通常可接受的规则以及中心极限定理对推断性统计学的重要性，30 常作为两个理论领域：**大样本统计**和**小样本统计**的分界线．如果 $n\geqslant 30$，则可以使用满足中心极限定理要求的**大样本方法**．如果 $n<30$，则使用第十四章和第十六章介绍的**小样本方法**．

例 13.15 在一所大学，四年级的 7300 个学生的 GPA(总平均成绩，见上册习题 6.43)分布有参数 $\mu=3.19$ 和 $\sigma=0.24$．如果无放回抽取一个包含 36 个学生的随机样本，试问：样本均值 GPA 与 μ 相差在 0.4σ 以内的概率是多少？

解 因为 $n\geqslant 30$ 且 $N\geqslant 2n$，可以使用从有限总体无放回抽样(见 13.17 节)的中心极限定理．由于 $n\leqslant 0.05N$，不必使用有限总体的修正因子．因此，可以认为 $\bar{X}$ 近似服从参数为 $\mu_{\bar{x}}=\mu=3.19$ 和 $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{0.24}{\sqrt{36}}=0.04$ 的正态分布．由这些信息，可以得到这个问题的近似结果．

$$\begin{aligned}
P(\mu-0.4\sigma_{\bar{x}}\leqslant \bar{X}\leqslant \mu+0.4\sigma_{\bar{x}})
&=P[3.19-(0.4\times 0.04)\leqslant \bar{X}\leqslant 3.19+(0.4\times 0.04)]\\
&=P(3.174\leqslant \bar{X}\leqslant 3.206)
\end{aligned}$$

由公式(13.14)的近似形式，

$$P(a\leqslant \bar{X}\leqslant b)\approx P\left(\frac{a-\mu_{\bar{x}}}{\sigma_{\bar{x}}}\leqslant Z\leqslant \frac{b-\mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)=P(z_a\leqslant Z\leqslant z_b) \tag{13.15}$$

须解

$$\begin{aligned}
P(3.174\leqslant \bar{X}\leqslant 3.206)&\approx P\left(\frac{3.174-3.19}{0.04}\leqslant Z\leqslant \frac{3.206-3.19}{0.04}\right)\\
&=P(-0.40\leqslant Z\leqslant 0.40)
\end{aligned}$$

对 $z=0.40$ 查表 A.5，

$$P(3.174\leqslant \bar{X}\leqslant 3.206)\approx 2(0.1554)=0.3108$$

13.19 样本和的抽样分布

容量为 n 的所有随机样本$(X_1,X_2,\cdots,X_n)$均取自无限大的正态分布总体(μ,σ^2,σ)，对每一样本求和(**样本和**，用 Y 表示)

$$Y=\sum_{i=1}^{n}X_i \tag{13.16}$$

Y 可以取值

$$y=\sum_{i=1}^{n}x_i \tag{13.17}$$

在数学上可以证明，在这些抽样条件下，**样本和的**(随机变量 Y 的)**抽样分布**精确服从参数为 $\mu_y=n\mu$，$\sigma_y^2=n\sigma^2$ 和 $\sigma_y=\sqrt{n}\sigma$(这里 σ_y 是**样本和的标准差**)的正态分布．

与任何正态分布的随机变量一样，如果 Y 服从正态分布，则 Z 变换

$$Z=\frac{Y-\mu_y}{\sigma_y} \tag{13.18}$$

将把正态分布变换为标准正态分布(见 12.7 节)．因此，由公式(12.7)

$$P(a\leqslant Y\leqslant b)=P\left(\frac{a-\mu_y}{\sigma_y}\leqslant Z\leqslant \frac{b-\mu_y}{\sigma_y}\right)=P(z_a\leqslant Z\leqslant z_b) \tag{13.19}$$

可以计算样本和的抽样分布的有关概率

例 13.16　一家渔业公司拥有 15 只船，凭多年的经验发现，捕捞量/船/天服从 $\mu=500\text{lb}$ 和 $\sigma=40\text{lb}$ 的正态分布．如果某天有 9 只船出海捕鱼，则该天捕捞量在 4400lb 和 4600lb 之间的概率是多少?

解 ☞　要求解的概率是：对日捕捞量正态总体，9 只船的日捕捞量的和 $y=\sum_{i=1}^{9}x_i$ 在 4400lb 和 4600lb 之间的概率．由于随机变量(日捕捞量 X)服从正态总体，故随机变量样本和 Y 也服从正态分布，并有 $\mu_y=n\mu=9\times500=4500\text{lb}$ 和 $\sigma_y=\sqrt{n}\sigma=\sqrt{9}\times40=120\text{lb}$．因此，由公式(13.19)得

$$P(4400\leqslant Y\leqslant4600)=P\left(\frac{4400-4500}{120}\leqslant Z\leqslant\frac{4600-4500}{120}\right)$$
$$=P(-0.83\leqslant Z\leqslant0.83)$$

对 $z=0.83$ 查表 A.5，

$$P(4400\leqslant \bar{Y}\leqslant4600)=2(0.2967)=0.5934$$

13.20　中心极限定理应用于样本和的抽样分布

无论对有限总体有放回抽样，还是对无限总体抽样，中心极限定理均可应用于样本和的理论抽样分布(见 13.19 节)：

在数学上可以证明，如果所有容量为 n 的随机样本，均来自有限总体的有放回抽样或无限总体，其中总体有有限参数(μ,σ^2,σ)，对每一样本计算样本和 $Y=y$，如果 n 足够大，则样本和 Y 的抽样分布近似服从参数为$\mu_y=n\mu$，$\sigma_y^2=n\sigma^2$ 和 $\sigma_y=\sqrt{n}\sigma$ 的正态分布．

同样，"足够大"意思是 $n\geqslant30$(见 13.18 节)．如果 Y 近似服从正态分布，用公式(13.18)对 Y 标准化，则 Z 的分布近似服从标准正态分布．公式(13.19)的以下形式是 Y 和 Z 分布之间的基本概率关系：

$$P(a\leqslant Y\leqslant b)\approx P\left(\frac{a-\mu_y}{\sigma_y}\leqslant Z\leqslant\frac{b-\mu_y}{\sigma_y}\right)=P(z_a\leqslant Z\leqslant z_b)\qquad(13.20)$$

稍做修改，中心极限定理可用于有限总体的无放回抽样．具体修改是(见 13.17 节)：N 至少是 n 的 2 倍($N\geqslant2n$)，而且，如果 $n\leqslant0.05N$ 不成立，σ_y 应该用以下公式计算：

$$\sigma_y=\sqrt{n}\sigma\times\sqrt{\frac{N-n}{N-1}}\qquad(13.21)$$

例 13.17　在一家大型计算机软件公司，上周 1000 个程序员的加班时间的分布有参数 $\mu=5.75\text{hr}$(小时)和 $\sigma=0.48\text{hr}$．如果得到 36 个程序员的一个随机样本(无放回)，则他们在上周的加班时间总和在 202hr 与 210hr 之间的概率是多少?

解 ☞　因为 $n\geqslant30$ 且 $N\geqslant2n$，由中心极限定理知，$Y=\sum_{i=1}^{36}X_i$ 近似服从正态分布，且参数 $\mu_y=n\mu=36\times5.75=207.0\text{hr}$，而且由于 $n\leqslant0.05N$，$\sigma_y=\sqrt{n}\sigma=\sqrt{36}\times0.48=2.88\text{hr}$．由这些信息，得到问题的近似解．

$$P(202\leqslant Y\leqslant210)$$

由公式(13.20)知

$$P(202\leqslant Y\leqslant210)\approx P\left(\frac{202-207.0}{2.88}\leqslant Z\leqslant\frac{210-207.0}{2.88}\right)$$
$$=P(-1.74\leqslant Z\leqslant1.04)$$

对 $z=-1.74$ 和 $z=1.04$ 查表 A.5，

$$P(202\leqslant Y\leqslant210)\approx0.4591+0.3508=0.8099$$

13.21　二项总体的抽样

二项总体的组成元素只能是两种可能类别之一：是/否，正/负，男性/女性，红色/非红色，缺

陷/无缺陷，等等. 一个测量总体，无论是离散型还是连续型，都能变换成为一个二项总体. 例如，成年男性身高的测量值总体，将每个值分为低于 67in(身高<67in)或等于大于 67in(身高≥67in)，就变换为一个二项总体.

n 重 Bernoulli 试验(见 11.2 节)的任一集合都可以看做是来自二项总体的一个容量为 n 的随机样本，其元素可分为成功或者失败. 并且，对任一 Bernoulli 试验，成功的概率 p 保持不变，等于二项总体的**成功比率**

$$p = \frac{\text{总体中的成功次数}}{\text{总体容量}}$$

(在一些统计学著作中，这个总体比率用 π 表示，即希腊字母 pi). 反之，在 Bernoulli 试验条件下，取自二项总体的一个容量为 n 的随机样本，可以看作是多重 Bernoulli 试验.

在二项总体中，如果对每个失败元素赋值 0，对每个成功元素赋值 1，则总体的相对频数分布可以用具有以下性质的二项概率分布(见 11.5 到 11.9 节)来很好的拟合：(1) 离散随机变量 X 表示在一次($n=1$)Bernoulli 试验中成功的次数($X=0$ 或 $X=1$)，(2)对一次试验的两个可能结果，概率函数 $f(x)$给出 $f(0)=1-p=q$ 和 $f(1)=p$，(3) 分布的参数是 $n=1$ 和 p，(4) 分布的均值[公式(11.4)]是 $\mu=np=1\times p=p$，(5) 分布的方差[公式(11.6)]是 $\sigma^2=npq=1\times p\times q=pq$，(6) 分布的标准差[公式(11.7)]是 $\sigma=\sqrt{\sigma^2}=\sqrt{pq}$.

成功比率 p 是二项总体的一个参数，可以用来自总体的一个样本的成功比率进行估计. 样本比率是样本估计量(见 13.8 节)，如下计算

$$\bar{P} = \frac{Y}{n} \tag{13.22}$$

这里，$\bar{P}$ 是随机变量，表示样本的成功比率(在一些著作中用 $\hat{P}$ 或 P' 表示)；Y 是二项随机变量，表示样本中的成功次数；n 是样本容量. (符号 Y 用来既表示成功次数又表示样本和的原因，见 13.24 节). 随机变量 $\bar{P}$ 可取实值 $\bar{P}=\bar{p}$. $\bar{P}$ 的概率分布和二项随机变量 Y 的概率分布相同，只是在尺度上有所改变. 因此，对于成功比率为 p 的二项总体，在 Bernoulli 试验条件下得到一个容量为 n 的样本，如果 a 是样本中的成功次数，则

$$P(Y=a) = P\left(\frac{Y}{n} = \frac{a}{n}\right) = P\left(\bar{P} = \frac{a}{n}\right) \tag{13.23}$$

例 13.18　有限二项总体的成功比率是 $p=0.4$，在 Bernoulli 试验条件下从该总体抽取一个随机样本($n=10$). 二项随机变量 Y，用于记录样本中的成功次数，能从 $Y=0$ 到 $Y=10$ 取值. 用二项概率函数[公式(11.2)]计算样本包含如下情形的概率：(a)3 次成功，(b)成功比率为 0.3，(c)30% 成功. 接着计算概率：(d)$\bar{P}$ 至多为 0.4，(e)$\bar{P}$ 至少为 0.3.

解　(a) 因为这个样本等价于 n 次二项试验的结果，用公式(11.2)的如下形式求解：

$$f(y) = \binom{n}{y} p^y q^{n-y},\ y = 0, 1, 2, \cdots, n$$

本例中

$$f(3) = \binom{10}{3}(0.4)^3(0.6)^7 = \frac{10!}{3!(10-3)!}(0.4)^3(0.6)^7$$

$$= (120)(0.064)(0.027994) = 0.214994$$

(b) 由公式(13.23)知，如果 $a=3$ 和 $n=10$，则

$$P(Y=3) = P\left(\frac{Y}{10} = \frac{3}{10}\right) = P(\bar{P} = 0.3) = 0.214994$$

(c) 比率可以用分数，小数或百分数表示，从而

$$P(P = 30\%) = P(\bar{P} = 0.3) = 0.214994$$

(d) 将公式(13.23)变换为不等式(见上册,1.23节)并代入 $a=4$ 和 $n=10$,

$$P(\bar{P} \leqslant 0.4) = P\left(\frac{Y}{10} \leqslant \frac{4}{10}\right) = P(Y \leqslant 4)$$

对 $a=4$, $n=10$ 和 $p=0.4$ 查附录表 A.3,

$$P(\bar{P} \leqslant 0.4) = P(Y \leqslant 4) = F(4) = 0.6331$$

(e) 将公式(13.23)变换为不等式并代入 $a=3$ 和 $n=10$,

$$P(\bar{P} \geqslant 0.3) = P\left(\frac{Y}{10} \geqslant \frac{3}{10}\right) = P(Y \geqslant 3)$$

由于 Y 只能取整数 $y=0, 1, 2, \cdots, 10$,

$$P(Y \geqslant 3) = 1 - P(Y \leqslant 2) = 1 - F(2)$$

对 $a=2$, $n=10$ 和 $p=0.4$ 查附录表 A.3,

$$P(\bar{P} \geqslant 0.3) = 1 - F(2) = 1 - 0.1673 = 0.8327$$

13.22　成功次数的抽样分布

在 Bernoulli 试验条件下从成功比率为 p 的二项总体抽取一个容量为 n 的随机样本, $Y=y$ 次成功的概率可以用二项概率函数[公式(11.2)]确定,其中 $y=0, 1, 2, \cdots, n$. 因此,如果容量为 n 的所有可能样本均来自 Bernoulli 试验条件下的一个有限二项总体,并且对每个样本计算出成功次数,则**成功次数的理论抽样分布(或计数抽样分布或发生次数的抽样分布)**服从二项概率分布(见 11.5 到 11.9 节),并有

$$E(Y) = \mu_y = np \tag{13.24}$$

$$\mathrm{var}(Y) = \sigma_y^2 = npq \tag{13.25}$$

$$\mathrm{st\ dev}(Y) = \sigma_y = \sqrt{npq} \tag{13.26}$$

这个标准差称为**成功次数的标准误**. 同先前的抽样分布一样(见 13.6 节),它与成功次数的经验抽样分布有联系,是容量为 n 的样本取实际值 $Y=y$ 的相对频数分布. 由于 p 在 Bernoulli 试验中保持不变,故上述公式对于无限总体抽样或者有限总体有放回抽样同样成立. 如果从有限总体无放回抽样,但 $n \leqslant 0.05N$,则:(1)可以假设 p"基本上"不变,并且(2)计算 σ_y^2 或 σ_y 不必使用有限总体的修正因子.

13.23　比率的抽样分布

如果容量为 n 的所有可能样本均来自 Bernoulli 试验条件下的一个无限二项总体,并且对每个样本计算出成功比率 $\bar{P}=\bar{p}$,则**比率的理论抽样分布**与成功次数的抽样分布(见 13.22 节)相同,只是在尺度上有所改变. 不过,这种尺度变化导致不同的均值,方差和标准差的公式:

$$E(\bar{P}) = \mu_{\bar{p}} = p \tag{13.27}$$

$$\mathrm{var}(\bar{P}) = \sigma_{\bar{p}}^2 = \frac{pq}{n} \tag{13.28}$$

$$\mathrm{st\ dev}(\bar{P}) = \sigma_{\bar{p}} = \sqrt{\frac{pq}{n}} \tag{13.29}$$

这个标准差称为**比率的标准误**. 同成功次数的抽样分布一样,公式可应用于无限二项总体的抽样,也可应用于有限二项总体的有放回抽样;如果 $n \leqslant 0.05N$,也可应用于有限总体的无放回抽样. 而且,同成功次数的抽样分布一样,比率的理论抽样分布与比率的经验抽样分布有联系.

例 13.19 如果 Y 是随机变量,a 和 b 是常数,且 $a>0$,则 $R=aY+b$ 一般也是随机变量.对于这样的两个变量,它们的均值,方差和标准差有如下关系:

$$E(R) = aE(Y) + b$$

$$\mathrm{var}(R) = a^2 \times \mathrm{var}(Y)$$

$$\mathrm{st\ dev}(R) = a \times \mathrm{st\ dev}(Y)$$

用上述关系说明:从成功次数的抽样分布(见 13.22 节)到比率的抽样分布,由于尺度变化而引起的均值,方差和标准差公式的变化.

解 ☞ 公式(13.22)表明

$$\bar{P} = \frac{Y}{n}$$

也可写成

$$\bar{P} = n^{-1}Y + 0$$

因此,在上述关系式中代入:$R=\bar{P}$, $Y=Y$, $a=n^{-1}$, $b=0$;并且由 13.22 节知,$E(Y)=np$, $\mathrm{var}(Y)=npq$, $\mathrm{st\ dev}(Y)=\sqrt{npq}$,

$$E(\bar{P}) = n^{-1}(np) + 0 = \frac{1}{n}np + 0 = p$$

$$\mathrm{var}(\bar{P}) = (n^{-1})^2(npq) = \left(\frac{1}{n}\right)\left(\frac{1}{n}\right)(npq) = \frac{pq}{n}$$

$$\mathrm{st\ dev}(\bar{P}) = n^{-1}\sqrt{npq} = \frac{\sqrt{npq}}{n} = \frac{\sqrt{npq}}{\sqrt{n^2}} = \sqrt{\frac{pq}{n}}$$

13.24 中心极限定理应用于成功次数的抽样分布

12.12 节指出,如果样本容量"足够大"($np\geqslant5$ 且 $nq\geqslant5$),则:(1)可以认为二项随机变量 Y 近似服从参数为 $\mu_y=np$, $\sigma_y^2=npq$, $\sigma_y=\sqrt{npq}$ 的正态分布.(2)如果 Y 用公式(13.18)进行过标准化,则得到的变量 Z 近似服从标准正态分布.现在,可以进一步指出:要对上述正态近似进行理论证明,只需将中心极限定理应用于样本和($Y=\sum_{i=1}^{n}X_i$)(见 13.20 节)即可;因为,样本的成功次数 Y,可以看做样本和的一个特例.

为理解以上结论,回顾简单随机样本的数学定义(见 13.3 节):n 个独立随机变量 X_1, $X_2,\cdots,X_n$,每一个都与抽样总体有相同的概率分布.这里,总体是二项总体,由一次试验($n=1$)中的成功($X=1$)和失败($X=0$)组成的二项分布($\mu=p$, $\sigma^2=pq$, $\sigma=\sqrt{pq}$).于是,样本中的每个独立随机变量 X_i 都有该二项分布,并且取值只能是 $X=0$ 和 $X=1$.因此,样本的成功次数 Y 实际上是样本统计量[公式(13.16)]:$Y=\sum_{i=1}^{n}X_i$ 的一个计算结果.并且,由应用于样本和的中心极限定理知,如果 n"足够大",则统计量 Y 的抽样分布近似服从正态分布.

13.25 中心极限定理应用于比率的抽样分布

13.23 节指出,比率的抽样分布与成功次数服从二项分布的抽样分布相同,但在尺度上有所改变.于是,必然成立:如果 n"足够大"($np\geqslant5$ 且 $nq\geqslant5$),比率的抽样分布近似服从正态分布,且有 $\mu_{\bar{p}}=p$, $\sigma_{\bar{p}}^2=\frac{pq}{n}$ 和 $\sigma_{\bar{p}}=\sqrt{\frac{pq}{n}}$;并且在上述条件下,

$$Z = \frac{\bar{P} - \mu_{\bar{p}}}{\sigma_{\bar{p}}} \tag{13.30}$$

将近似服从标准正态分布.以上结果的理论证明也缘于中心极限定理,但此时是用于样本均值

(见 13.15 到 13.17 节).这样,$\bar{P}$ 被看做样本均值的一个特例.

为理解这一点,由前知 $\bar{P}=\frac{Y}{n}$[公式(13.22)]和 $Y=\sum_{i=1}^{n} X_i$(见 13.24 节),从而

$$\bar{P}=\frac{\sum_{i=1}^{n} X_i}{n} \tag{13.31}$$

这样 $\bar{P}$ 实际上是每个样本(或每次试验)的平均成功次数.

由于 $\bar{P}$ 实际上是样本均值,故可以从 $\mu_{\bar{x}}$ 导出 $\mu_{\bar{p}}$,从 $\sigma_{\bar{x}}^2$ 导出 $\sigma_{\bar{p}}^2$,从 $\sigma_{\bar{x}}$ 导出 $\sigma_{\bar{p}}$.于是,如果二项总体无限或者有限但有放回抽样或者 $n \leqslant 0.05N$ 且有限但无放回抽样,则有

$$\mu_{\bar{x}}=\mu=\mu_{\bar{p}}$$

$$\sigma_{\bar{x}}^2=\frac{\sigma^2}{n}=\sigma_{\bar{p}}^2$$

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\sigma_{\bar{p}}$$

对于二项总体(见 13.21 节),已知 $\mu=p$, $\sigma^2=pq$ 和 $\sigma=\sqrt{pq}$,因此

$$\mu_{\bar{p}}=p$$

$$\sigma_{\bar{p}}^2=\frac{pq}{n}$$

$$\sigma_{\bar{p}}=\frac{\sqrt{pq}}{\sqrt{n}}=\sqrt{\frac{pq}{n}}$$

这是 13.23 节给出的公式并在例 13.19 有不同的推导.

注:本节指出,如果 $np \geqslant 5$ 且 $nq \geqslant 5$*,则可以使用中心极限定理;但是,在这之前(13.18 节),使用中心极限定理的要求是 $n \geqslant 30$.为什么存在这种差异呢?如果将中心极限定理应用于服从二项分布的抽样分布,则使用规则 $np \geqslant 5$ 且 $nq \geqslant 5$,因为随着 p 越来越接近 0.5(见 12.12 节),这些分布变得越来越对称,于是,如果 $p=0.5$,一个大小为 $n=10$ 的样本即允许使用中心极限定理.然而,对于该规则,随着 p 远离 0.5,必须增大 n.因此,当 $p=0.1$ 时,n 必须至少为 50.

13.26　用比率的抽样分布的正态近似计算概率

如果在 Bernoulli 试验条件下样本取自一个二项总体,则比率的抽样分布服从二项分布且有 $\mu_{\bar{p}}=p$, $\sigma_{\bar{p}}^2=\frac{pq}{n}$ 和 $\sigma_{\bar{p}}=\sqrt{\frac{pq}{n}}$.如果 $np \geqslant 5$ 且 $nq \geqslant 5$,则可以认为该分布近似服从正态分布(见 13.25 节),并且公式(13.30)$Z=\frac{\bar{P}-\mu_{\bar{p}}}{\sigma_{\bar{p}}}$ 近似服从标准正态分布.这样,似乎可以由关系式

$$P(a \leqslant \bar{P} \leqslant b) \approx P\left(\frac{a-\mu_{\bar{p}}}{\sigma_{\bar{p}}} \leqslant Z \leqslant \frac{b-\mu_{\bar{p}}}{\sigma_{\bar{p}}}\right)=P(z_a \leqslant Z \leqslant z_b) \tag{13.32}$$

计算概率,这也是在许多统计学著作中可以找到的解法.然而,随机变量 $\bar{P}$ 是一个离散随机变量;一些著作建议,同二项分布的正态近似[见例 12.10(b)]一样,应该进行**连续性修正**,尤其是对小样本.建议使用的修正是适当地加上或减去 $1/2n$.我们将在下面的例子以及本章后面的合适之处用到这种修正.不过,由于这个修正因子很少使用,也为了便于解释,在后面几章关

* 此处原文有误.——译者注

于比率的估计理论(第 15 章)或假设检验理论(第 16 章)中,我们将不使用该因子.

例 13.20　选举前的一次民意测验表明,一座城市的 100000 个注册选民总体的大部分($p=0.51$)支持候选人 A 当选市长. 然而,抽取 100 个选民的民意测验表,发现这个随机样本的比率即支持 A 的比率是 $\bar{P}=0.49$. 假设选民总体中支持 A 的比率确实是 $p=0.51$,则 $P(0.40\leqslant\bar{P}\leqslant0.64)$是多少?

解 ☞　如果认为这是在 Bernoulli 试验条件下对二项总体的抽样,则支持 A 的比率的抽样分布服从二项分布且有 $\mu_{\bar{p}}=p=0.51$, $\sigma_{\bar{p}}^2=\frac{pq}{n}=\frac{0.51\times0.49}{100}=0.002499$ 和 $\sigma_{\bar{p}}=\sqrt{\frac{pq}{n}}=\sqrt{0.002499}=0.049989999\approx0.05$. 由于 $np=51$ 和 $nq=49$,可以认为该分布是近似正态的. 为使用 Z 变换将分布变换为标准正态分布,首先应对公式(13.32)使用连续性修正因子 $1/2n$:

$$P\left(0.40-\frac{1}{2n}\leqslant\bar{P}\leqslant0.60+\frac{1}{2n}\right)$$

$$\approx P\left(\frac{0.40-1/2n-0.51}{0.05}\leqslant Z\leqslant\frac{0.60+1/2n-0.51}{0.05}\right)$$

将修正因子值 $\frac{1}{2n}=\frac{1}{2(100)}=0.005$ 代入公式,

$$P(0.395\leqslant\bar{P}\leqslant0.605)\approx P\left(\frac{0.395-0.51}{0.05}\leqslant Z\leqslant\frac{0.605-0.51}{0.05}\right)$$

$$=P(-2.30\leqslant\bar{P}\leqslant1.90)$$

对 $z=2.30$ 和 $z=1.90$ 查表 A.5,

$$P(0.40\leqslant\bar{P}\leqslant0.60)\approx0.4893+0.4713=0.9606$$

习 题 解 答

均值的抽样分布

13.1　对上册附录表 A.1 进行有放回抽样(接受重复数字和重复样本),采用例 13.3 从 5 个数字 0,1,2,3,4 的总体中抽出 100 个 2 个数字($n=2$)随机样本的方法. 抽样结果列在表 13.7 中,对于 25 个可能样本,该表还给出每一个样本的选取频数和样本均值. 由上述信息,用表格形式给出这个概率问题的均值的经验抽样分布.

表 13.7

[(样本)$\bar{x}_i$]	频数 f_i	[(样本)$\bar{x}_i$]	频数 f_i	[(样本)$\bar{x}_i$]	频数 f_i
[(0, 0)0.0]	5	[(1, 3)2.0]	4	[(3, 1)2.0]	5
[(0, 1)0.5]	6	[(1, 4)2.5]	5	[(3, 2)2.5]	5
[(0, 2)1.0]	4	[(2, 0)1.0]	3	[(3, 3)3.0]	2
[(0, 3)1.5]	3	[(2, 1)1.5]	6	[(3, 4)3.5]	2
[(0, 4)2.0]	2	[(2, 2)2.0]	4	[(4, 0)2.0]	3
[(1, 0)0.5]	7	[(2, 3)2.5]	5	[(4, 1)2.5]	1
[(1, 1)1.0]	4	[(2, 4)3.0]	4	[(4, 2)3.0]	7
[(1, 2)1.5]	3	[(3, 0)1.5]	1	[(4, 3)3.5]	5
				[(4,4)4.0]	4

解 ☞　均值的经验抽样分布是一个相对频数分布. 表 13.8 给出该数据集的经验抽样分布. 分布的相对频数值近似等于表 13.2 所示的均值的理论抽样分布的概率值. 但是由于样本数(100)与抽样总体(∞)相比很小,近似不很精确.

13.2　总体是三个($n=3$)数字:7,10,12,使用有放回抽样得到总体的容量为 $n=2$ 的随机样本,给出均值的离散型抽样分布. 要求:(a)列出来自总体的所有可能的样本,以及每一样本的算术平均值.(b)用频数和相对频数分布概括这些算术平均值.(c)计算每一均值的抽取概率,并用概率表描述得到的抽样分布.

表 13.8

样本均值 $\bar{x}_i$	相对频数 f_i/n_i
0.0	0.05
0.5	0.13
1.0	0.11
1.5	0.13
2.0	0.18
2.5	0.16
3.0	0.13
3.5	0.07
4.0	0.04
$\sum$	1.00

解 (a)由计数规则:乘法原理(见上册,9.12 节)知,有($n_1 \times n_2 = 3 \times 3 = 9$)个容量为 $n=2$ 的可能随机样本.这些样本及其均值[(样本)$\bar{x}$]是

$$[(7,\ 7)7.0]\quad [(10,\ 7)8.5]\quad [(12,\ 7)9.5]$$

$$[(7,\ 10)8.5]\quad [(10,\ 10)10.0]\quad [(12,\ 10)11.0]$$

$$[(7,\ 12)9.5]\quad [(10,\ 12)11.0]\quad [(12,\ 12)12.0]$$

表 13.9

样本均值 $\bar{x}_i$	频数 f_i	相对频数 $f_i/N_{\bar{x}}$
7.0	1	0.111111
8.5	2	0.222222
9.5	2	0.222222
10.0	1	0.111111
11.0	2	0.222222
12.0	1	0.111111
$\sum$	9	0.999999

(b) 要求的频数和相对频数在表 13.9 给出,这里 $N_{\bar{x}}$ 表示可能的 $\bar{x}$ 值的个数.

(c) 由于 9 个可能样本中每一个的抽取概率均为 $\frac{1}{9}$,故使用例 13.4(c)的方法,可得到理论抽样分布,如表 13.10 所示.

表 13.10

样本均值 $\bar{x}$	概率 $f(\bar{x})$
7.0	0.111111
8.5	0.222222
9.5	0.222222
10.0	0.111111
11.0	0.222222
12.0	0.111111
$\sum$	0.999999

13.3 总体是三个数字:7,10,12,使用无放回抽样得到总体的容量为 $n=2$ 的随机样本,给出均值的离散型抽样分布.要求:(a)列出来自总体的所有可能的样本以及每一样本的算术平均值.(b)用频数和相对频数分布概括这些算术平均值,(c)计算每一均值的抽取概率,并用概率表描述得到的抽样分布.

解 (a) 与例 13.5(a)一样,可能的随机样本的总数用计数规则:组合[上册,公式(9.12)]计算,

$$_3C_2 = \binom{3}{2} = \frac{3!}{2!(3-2)!} = 3$$

这些样本及其均值[(样本)$\bar{x}$]是

[(7, 10)8.5]　[(7, 12)9.5]　[(10, 12)11.0]

(b) 要求的频数和相对频数在表 13.11 给出,这里 $N_{\bar{x}}$ 表示可能的 $\bar{x}$ 值的个数.

表 13.11

样本均值 $\bar{x}_i$	频数 f_i	相对频数 $f_i/N_{\bar{x}}$
8.5	1	0.333333
9.5	1	0.333333
11.0	1	0.333333
$\sum$	3	0.999999

(c) 由于 9 个可能样本中每一个的抽取概率均为 $\frac{1}{3}$,使用例 13.4(c)的方法,可得到理论抽样分布,如表 13.12 所示.

表 13.12

样本均值 $\bar{x}$	概率 $f(\bar{x})$
8.5	0.333333
9.5	0.333333
11.0	0.333333
$\sum$	0.999999

13.4　城镇的某一地区建有 6 座待售住房,分别有卧室 1 间,2 间,3 间,4 间,5 间和 6 间.一个潜在的购房者随机选择两座不同房屋进行参观(即无放回抽样),试问:(a)不同样本的总个数是多少?(b)在一个样本中,平均有 4 间卧室的概率是多少?

解 (a) 与例 13.5(a)一样,可能样本的总个数用计数规则:组合[上册,公式(9.12)]计算,

$$_6C_2 = \binom{6}{2} = \frac{6!}{2!(6-2)!} = 15$$

(b) 所有的可能随机样本及其均值[(样本)$\bar{x}$]如下:

[(1, 2)1.5]

[(1, 3)2.0]　[(2, 3)2.5]

[(1, 4)2.5]　[(2, 4)3.0]　[(3, 4)3.5]

[(1, 5)3.0]　[(2, 5)3.5]　[(3, 5)4.0]　[(4, 5)4.5]

[(1, 6)3.5]　[(2, 6)4.0]　[(3, 6)4.5]　[(4, 6)5.0]　[(5, 6)5.5]

以上 15 个均值中有 2 个等于 4.0 间卧室.因此,得到平均有 4.0 间卧室的一个样本的概率是 2/15 或 0.1333.

均值的抽样分布的均值,方差和标准差

13.5　对表 13.10 的均值的理论抽样分布,计算均值的均值,方差和标准差.

解 由公式(13.2),

$$\mu_{\bar{x}} = \sum_{\bar{x}} \bar{x} f(\bar{x})$$

$$= 7.0\left(\frac{1}{9}\right) + 8.5\left(\frac{2}{9}\right) + 9.5\left(\frac{2}{9}\right) + 10.0\left(\frac{1}{9}\right) + 11.0\left(\frac{2}{9}\right) + 12.0\left(\frac{1}{9}\right)$$

$$= 9.6667 \text{ 或 } 9.67$$

由公式(13.4),

$$\sigma_{\bar{x}}^2=\sum_{\bar{x}}(\bar{x}-\mu_{\bar{x}})^2 f(\bar{x})$$

$$=(7.0-9.67)^2\left(\frac{1}{9}\right)+(8.5-9.67)^2\left(\frac{2}{9}\right)+(9.5-9.67)^2\left(\frac{2}{9}\right)$$

$$+(10.0-9.67)^2\left(\frac{1}{9}\right)+(11.0-9.67)^2\left(\frac{2}{9}\right)+(12.0-9.67)^2\left(\frac{1}{9}\right)$$

$$=2.1111 \text{ 或 } 2.111$$

由公式(13.6),

$$\sigma_{\bar{x}}=\sqrt{\sigma_{\bar{x}}^2}=\sqrt{2.1111}=1.4530 \text{ 或 } 1.45$$

13.6　对表 13.12 的均值的理论抽样分布,计算均值的均值,方差和标准差.

解　由公式(13.2),

$$\mu_{\bar{x}}=\sum_{\bar{x}}\bar{x}f(\bar{x})=8.5\left(\frac{1}{3}\right)+9.5\left(\frac{1}{3}\right)+11.0\left(\frac{1}{3}\right)$$

$$=9.6667 \text{ 或 } 9.67$$

由公式(13.4),

$$\sigma_{\bar{x}}^2=\sum_{\bar{x}}(\bar{x}-\mu_{\bar{x}})^2 f(\bar{x})$$

$$=(8.5-9.67)^2\left(\frac{1}{3}\right)+(9.5-9.67)^2\left(\frac{1}{3}\right)+(11.0-9.67)^2\left(\frac{1}{3}\right)$$

$$=1.0556 \text{ 或 } 1.056$$

由公式(13.6),

$$\sigma_{\bar{x}}=\sqrt{\sigma_{\bar{x}}^2}=\sqrt{1.0556}=1.0274 \text{ 或 } 1.03$$

13.7　一个有限总体,$N=60$ 且 $\sigma=2.0$,从中无放回抽取一个容量为 n 的随机样本,试说明:如果样本容量从 $n=4$ 增加到 $n=16$,$\bar{X}$ 作为 μ 的估计将更精密.

解　与例 13.11 一样,我们通过 $\sigma_{\bar{x}}$ 的变化来表明精密度随样本容量的变化.由于是从有限总体无放回抽样并且两个 n 值都大于 N 的 5% (见 13.10 节),故使用有限总体的修正因子,将 $N=60$, $\sigma=2.0$ 和 $n=4$ 代入公式(13.8):

$$\sigma_{\bar{x}}=\frac{2.0}{\sqrt{4}}\sqrt{\frac{60-4}{60-1}}=1.0\sqrt{0.949153}=0.974245 \text{ 或 } 0.97$$

将 n 从 4 增加到 16,

$$\sigma_{\bar{x}}=\frac{2.0}{\sqrt{16}}\sqrt{\frac{60-16}{60-1}}=0.5\sqrt{0.745763}=0.431788 \text{ 或 } 0.43$$

于是,样本容量增加 4 倍而 σ 保持不变,则 $\sigma_{\bar{x}}$ 减少 55% 多.

13.8　对表 13.8 中的均值的经验抽样分布,计算样本均值的均值(用 $\hat{\mu}_{\bar{x}}$ 表示,读作 mu-x-bar-hat)和样本均值的标准差(用 $\hat{\sigma}_{\bar{x}}$ 表示,读作 sigma-x-bar-hat).为完成计算,从表 13.8 消去相对频数列,代之以下 4 列:频数(f_i),频数×均值($f_i\bar{x}_i$),均值平方($\bar{x}_i^2$)和频数×均值平方($f_i\bar{x}_i^2$).然后,用上册公式(6.10)和(7.33)计算:

$$\hat{\mu}_{\bar{x}}=\frac{\sum_{i=1}^{k}f_i\bar{x}_i}{n_{\bar{x}}} \tag{13.33}$$

和

$$\hat{\sigma}_{\bar{x}}=\sqrt{\frac{\sum_{i=1}^{k}f_i\bar{x}_i^2-\frac{\left(\sum_{i=1}^{k}f_i\bar{x}_i\right)^2}{n_{\bar{x}}}}{n_{\bar{x}}-1}} \tag{13.34}$$

这里 $n_{\bar{x}}$ 表示样本数.

解 ☞ 要求的计算 $\hat{\mu}_{\bar{x}}$ 和 $\hat{\sigma}_{\bar{x}}$ 如表 13.13 所示. 可以看到,仅仅列出了无穷多个可能样本中的 100 个,并且计算值仅仅是 $\mu_{\bar{x}}$ [$\hat{\mu}_{\bar{x}}=1.92$ 估计 $\mu_{\bar{x}}=2.0$;见例 13.6(a)]和 $\sigma_{\bar{x}}$ [$\hat{\sigma}_{\bar{x}}=1.06$ 估计 $\sigma_{\bar{x}}=1.00$; 见例 13.9]的适当估计.

表 13.13

样本均值 $\bar{x}_i$	频数 f_i	$f_i\bar{x}_i$	$\bar{x}_i^2$	$f_i\bar{x}_i^2$
0.0	5	0.0	0.00	0.00
0.5	13	6.5	0.25	3.25
1.0	11	11.0	1.00	1.00
1.5	13	19.5	2.25	29.25
2.0	18	36.0	4.00	72.00
2.5	16	40.0	6.25	100.00
3.0	13	39.0	9.00	117.00
3.5	7	24.5	12.25	85.75
4.0	4	16.0	16.00	64.00
$\sum$	100	192.5		482.25

$$\hat{\mu}_{\bar{x}}=\frac{\sum f_i\bar{x}_i}{n_{\bar{x}}}=\frac{192.5}{100}=1.92$$

$$\hat{\sigma}_{\bar{x}}=\sqrt{\frac{\sum f_i\bar{x}_i^2-\frac{(\sum f_i\bar{x}_i)^2}{n_{\bar{x}}}}{n_{\bar{x}}-1}}=\sqrt{\frac{482.25-\frac{(192.5)^2}{100}}{99}}=1.0621,\quad \text{or} 1.06$$

用均值的离散型抽样分布计算概率

13.9 对于例 13.3 的 5 个数字总体:0,1,2,3,4 和表 13.2 给出的均值的离散型抽样分布,无放回抽取一个容量为($n=2$)的随机样本,计算样本均值($\bar{X}=\bar{x}$)大于 0.5 但至多为 2.0 的概率.

解 ☞ 上册 10.7 节离散变量的累积分布函数的一般规则也适用于离散变量 $\bar{X}$. 于是,由上册公式(10.4),

$$P(0.5<\bar{X}\leqslant 2.0)=F(2.0)-F(0.5)$$

从例 13.12 知,对该分布有 $F(2.0)=0.60$. 计算 $F(0.5)$,

$$F(0.5)=f(0.0)+f(0.5)=0.04+0.08=0.12$$

因此,

$$P(0.5<\bar{X}\leqslant 2.0)=0.60-0.12=0.48$$

13.10 对于例 13.3 的 5 个数字总体:0,1,2,3,4 和表 13.4 给出的均值的离散型抽样分布,无放回抽取一个容量为($n=2$)的随机样本,计算样本均值($\bar{X}=\bar{x}$)为 0.5,1.0 或 1.5 的概率.

解 ☞ 由公式(13.10)和取自表 13.4 的值

$$F(1.5)=\sum_{\bar{x}\leqslant 1.5}f(\bar{x})$$

$$=f(0.5)+f(1.0)+f(1.5)=0.1+0.1+0.2=0.4$$

13.11 对于例 13.3 的 5 个数字总体:0,1,2,3,4 和表 13.4 给出的均值的离散型抽样分布,无放回抽取一个容量为($n=2$)的随机样本,计算样本均值($\bar{X}=\bar{x}$)至少为 1.5 但小于 4.5 的概率.

解 ☞　对离散变量 $\bar{X}$ 由上册公式(10.7),

$$P(1.5 \leqslant \bar{X} < 4.5) = F(4.5) - F(1.5) + f(1.5) - f(4.5)$$

由于

$$F(4.5) = F(3.5) - \sum_{\bar{x} \leqslant 3.5} f(\bar{x}) = 1.0$$

$$F(1.5) = 0.4$$

$$f(1.5) = 0.2$$

$$f(4.5) = 0.0$$

因此,

$$P(1.5 \leqslant \bar{X} < 4.5) = 1.0 - 0.4 + 0.2 - 0.0 = 0.8$$

13.12　对于例 13.3 的 3 数字总体:3,4,5 和表 13.5 给出的均值的离散型抽样分布,无放回抽取一个容量为 $n=2$ 的随机样本,计算样本均值($\bar{X}=\bar{x}$)小于 5.0 但大于 3.0 的概率.

解 ☞　对离散变量 $\bar{X}$ 由上册公式(10.6)并利用表 13.5 的值,

$$P(3.0 < \bar{X} < 5.0) = F(5.0) - F(3.0) + f(5.0)$$
$$= 1.0 - 0.111111 - 0.111111 = 0.7778$$

13.13　对表 13.5 中均值的离散型抽样分布,从例 13.14(c)知,$\mu_{\bar{x}} = 4.0$ 和 $\sigma_{\bar{x}} = 0.577350 \approx 0.58$. 由以上信息,随机变量 $\bar{X}$ 在区间 $\mu_{\bar{x}} = 1.5\sigma_{\bar{x}}$ 上取值的概率是多少?

解 ☞　同上册例 10.12 的离散随机变量 X 一样,对于变量 $\bar{X}$,这个问题既有由 Chebyshev 定理得到的近似解,也有精确解. 应用于均值的抽样分布的 Chebyshev 定理是

对任意 $k \geqslant 1$,如果随机变量 $\bar{X}$ 有均值 $\mu_{\bar{x}}$ 和标准差 $\sigma_{\bar{x}}$,则 $\bar{X}$ 在区间 $\mu_{\bar{x}} \pm k\sigma_{\bar{x}}$ 上取值的概率至少是 $1-\frac{1}{k^2}$.

因此,在这里,概率至少是

$$1 - \frac{1}{k^2} = 1 - \frac{1}{1.5^2} = 1 - \frac{1}{2.25} = 1 - 0.444444 = 0.555556 \text{ 或 } 0.56$$

精确解要求对 $\bar{X}$ 在区间 $\mu_{\bar{x}} \pm 1.5\sigma_{\bar{x}}$ 上的所有可能取值的概率进行求和:

$$\mu_{\bar{x}} \pm 1.5\sigma_{\bar{x}} = 4.0 \pm 1.5(0.577350) \text{ 或 } 4.0 \pm 0.866025 \text{ 或从 } 3.13 \text{ 到 } 4.87$$

由于 $\bar{X}$ 在区间上只能取值 3.5,4.0 和 4.5,由表 13.5 可得

$$P(-1.5\sigma_{\bar{x}} \leqslant \bar{X} < 1.5\sigma_{\bar{x}}) = f(3.5) + f(4.0) + f(4.5)$$
$$= 0.222222 + 0.333333 + 0.222222$$
$$= 0.777777 \text{ 或 } 0.78$$

用均值的正态抽样分布计算概率

13.14　随机样本($n=9$)取自正态分布总体($\mu = 0.75, \sigma = 0.15$). (a)连续随机变量 $\bar{X}$ 的分布特征(形状,参数)是什么? (b)$\bar{X}$ 的取值 $\bar{x}$ 至多为 0.8 的概率是多少? (c)$\bar{X}$ 的取值 $\bar{x}$ 至少为 0.8 的概率是多少?

解 ☞　(a)从 13.14 节知 $\bar{X}$ 服从正态分布,且 $\mu_{\bar{x}} = \mu = 0.75$ 和 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.15}{\sqrt{9}} = 0.05$.

(b) 由公式(13.14)知

$$P(\bar{X} \leqslant 0.80) = P(-\infty < Z \leqslant 0.00) + P\left(0.00 \leqslant Z \leqslant \frac{0.80 - 0.75}{0.05}\right)$$
$$= 0.5 + P(0.00 \leqslant Z \leqslant 1.00)$$

因此,对 $z = 1.00$ 查表 A.5,

$$P(\bar{X} \leqslant 0.08) = 0.5 + 0.3413 = 0.8413$$

(c) 从上册公式(10.1)知,对连续型概率分布,

$$P(a < \bar{X} \leqslant b) = P(a \leqslant \bar{X} \leqslant b) = P(a < \bar{X} < b) = P(a \leqslant \bar{X} < b)$$

因此,

$$P(\bar{X} \geqslant 0.80) = 1 - P(\bar{X} \leqslant 0.80) = 1 - 0.8413 = 0.1587$$

13.15 随机样本($n=64$)取自正态分布总体($\sigma=0.15$). 如果 $P(X \leqslant 62.0)=0.9452$, 则 $\mu_{\bar{x}}$ 是多少?

解 ☞ 由公式(13.14)并由习题 12.11 的方法,可知

$$P(\bar{X} \leqslant 62.0) = P(Z \leqslant z_\gamma) = 0.9452$$

并且

$$P(Z \leqslant z_\gamma) = 0.9452 = P(-\infty < Z \leqslant 0) + P(0 \leqslant Z \leqslant z_\gamma)$$
$$= 0.5 + P(0 \leqslant Z \leqslant z_\gamma)$$

因此,

$$P(0 \leqslant Z \leqslant z_\gamma) = 0.9452 - 0.5 = 0.4452$$

查表 A.5 得

$$z_\gamma = 1.60$$

对公式(13.11)重新整理,

$$\bar{X} = \mu_{\bar{x}} + Z\sigma_{\bar{x}} \tag{13.35}$$

在公式(13.35)代入 $\bar{X}=\bar{x}=62.0$ 和 $\sigma_{\bar{x}}=\frac{10.0}{\sqrt{64}}=1.25$,

$$62.0 = \mu_{\bar{x}} + (1.60 \times 1.25) = \mu_{\bar{x}} + 2.0$$

从而

$$\mu_{\bar{x}} = 60.0$$

13.16 如果正态分布的随机变量 $\bar{X}$ 有 $\mu_{\bar{x}}=5.5$ 和 $\sigma_{\bar{x}}=0.45$,则 z_α 和 $\bar{x}_\alpha$ 是多少? 设 $\alpha=0.025$.

解 ☞ 从 12.10 节[公式(12.9)]知,当使用这些符号时,α 表示概率分布上尾的面积. 于是,对正态分布的连续随机变量 $\bar{X}$,

$$P(\bar{X} > \bar{x}_\alpha) = P\left(Z > \frac{\bar{x}_\alpha - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) = P(Z > z_\alpha) = \alpha \tag{13.36}$$

对于标准正态分布,它是图 12-9 的阴影面积. 本题中,有

$$P(Z > z_\alpha) = \alpha = 0.025$$

因此

$$P(0 \leqslant Z \leqslant z_{0.025}) = 0.5 - 0.025 = 0.475$$

查表 A.5,得到 $z_{0.025}=1.96$. 由公式(13.35),

$$\bar{x}_\alpha = \mu_{\bar{x}} + z_\alpha\sigma_{\bar{x}}$$
$$= 5.5 + (1.96 \times 0.45) = 5.5 + 0.882 = 6.382 \text{ 或 } 6.4$$

13.17 对硬质黏土上种植的胡萝卜(见习题 12.8 到 12.12),假设胡萝卜长度 X 服从 $\mu=11.5$cm和 $\sigma=1.15$cm 的正态分布. 对于 25 个这种胡萝卜的一个随机样本,样本均值 $\bar{X}$ 与 μ 相差在 0.5cm 之内的概率是多少?

解 ☞ 这个问题即是需要计算概率

$$P(\mu - 0.5\text{cm} \leqslant X \leqslant \mu + 0.5\text{cm})$$

已知 $\mu_{\bar{x}} = \mu = 11.5\text{cm}$ 和 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.15}{\sqrt{25}} = 0.23\text{cm}$. 于是，由公式(13.14)，

$$P(11.5 - 0.5 \leqslant \bar{X} \leqslant 11.5 + 0.5)$$

$$= P\left(\frac{11.0 - 11.5}{0.23} \leqslant Z \leqslant \frac{12.0 - 11.5}{0.23}\right)$$

$$P(11.0 \leqslant \bar{X} \leqslant 12.0) = P(-2.17 \leqslant Z \leqslant 2.17)$$

对 $z = 2.17$ 查表 A.5，

$$P(11.0\text{cm} \leqslant \bar{X} \leqslant 12.0\text{cm}) = 2(0.4850) = 0.9700$$

13.18 对习题 13.17 的胡罗卜，如果抽取容量为 49 的一个样本用于估计 μ，则估计误差至多为 0.25cm 的概率是多少？

解 本题中的误差即抽样误差 $\bar{x} - \mu$(见 13.11 节). 因此，必须计算出容量为 $n = 49$ 的样本的均值 $\bar{x}$ 在 $\mu_{\bar{x}} = \mu = 11.5\text{cm}$ 的 $\pm 0.25\text{cm}$ 范围之内取值的概率. 由于 $n = 49$，得到 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.15}{\sqrt{49}} = 0.164286\text{cm}$. 于是，由公式(13.14)

$$P(11.5 - 0.25 \leqslant \bar{X} \leqslant 11.5 + 0.25)$$

$$= P\left(\frac{11.25 - 11.5}{0.164286} \leqslant Z \leqslant \frac{11.75 - 11.5}{0.164286}\right)$$

$$P(11.25 \leqslant \bar{X} \leqslant 11.75) = P(-1.52 \leqslant Z \leqslant 1.52)$$

对 $z = 1.52$ 查表 A.5，

$$P(11.25\text{cm} \leqslant \bar{X} \leqslant 11.75\text{cm}) = 2(0.4357) = 0.8714$$

13.19 某文化人类学家使用放射性测时技术发现，一种骨质工具的年龄 X 基本上服从正态分布($\mu = 12500$ 年，$\sigma = 400$ 年). 如果从数千个这种工具中抽取 150 个随机样本，其中每个样本包含 $n = 4$ 个工具，并分别测定时间，则有多少个样本的均值为 12800 年或更长？

解 首先计算出 $P(\bar{X} \geqslant 12800)$. 已知 $\mu_{\bar{x}} = \mu = 12500$ 年和 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{400}{\sqrt{4}} = 200$ 年. 因此，

$$P(\bar{X} \geqslant 12800) = P\left(Z \geqslant \frac{12800 - 12500}{200}\right) = P(Z \geqslant 1.50)$$

对 $z = 1.50$ 查表 A.5，

$$P(Z \geqslant 1.5) = 0.5 - 0.4332 = 0.0668$$

这样，150 个样本中均值为 12800 年或更长的期望个数是

$$0.0668 \times 150 = 10.02 \text{ 或 } 10.0$$

中心极限定理

13.20 对例 13.15 的 GPA 分布($\mu = 3.19$，$\sigma = 0.24$)，已知一个无放回抽样得到的包含 36 个四年级学生的随机样本，有 $\mu_{\bar{x}} = \mu = 3.19$，$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 0.04$. 对这类 $n = 36$ 的样本，如果 $P(|\bar{X} - \mu| \leqslant a) = 0.34$，则 a 是多少？

解 问题是需要解出公式

$$P(\mu - a \leqslant \bar{X} \leqslant \mu + a) = 0.34$$

中的 a 值. 对于有限总体无放回抽样问题，由于 $n \geqslant 30$ 且 $N \geqslant 2n$，故可以使用中心极限定理(见 13.17 节)，由公式(13.15)：

$$P(\mu - a \leqslant \bar{X} \leqslant \mu + a) = 0.34 \approx P\left(\frac{\mu - a - (\mu_{\bar{x}} = \mu)}{\sigma_{\bar{x}}} \leqslant Z \leqslant \frac{\mu + a - (\mu_{\bar{x}} = \mu)}{\sigma_{\bar{x}}}\right)$$

$$\approx P\left(\frac{-a}{\sigma_{\bar{x}}} \leqslant Z \leqslant \frac{a}{\sigma_{\bar{x}}}\right)$$

对概率$\frac{0.34}{2}=0.17$,表A.5中最接近的z值恰好是0.44.因此,

$$\frac{a}{\sigma_{\bar{x}}} \approx 0.44$$

已知$\sigma_{\bar{x}}=0.04$,

$$a \approx 0.44 \times 0.04 = 0.0176$$

13.21 一个化妆品制造商雇有1500个入户推销员,上月他们的平均销售额是$\mu=\$3100$,标准差$\sigma=\$350$.如果无放回地抽取49个推销员作为随机样本,则该样本在上月的平均销售额少于\$3000的概率是多少?

解 ☞ 由于$n \geqslant 30$且$N \geqslant 2n$,可以使用从有限总体无放回抽样(见13.17节)的中心极限定理,即可以认为$\bar{X}$的分布是近似正态的,有参数$\mu_{\bar{x}}=\mu=\$3100$,并且因为$n \leqslant 0.05N$,$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{350}{\sqrt{49}}=\$50$.由上述信息,这个问题的近似解可表述为

$$P(\bar{X} < \$3000)$$

由公式(13.15)知

$$P(\bar{X} < \$3000) \approx P\left(Z < \frac{\$3000-\$3100}{\$50}\right) = P(Z < -2.00)$$

对$z=2.00$查表A.5,

$$P(\bar{X} < \$3000) \approx 0.5 - 0.4772 = 0.0228$$

13.22 用于牙科的一种麻醉剂从注射到消除牙床注射区域的敏感性平均需要$\mu=290\text{sec}$($\sigma=25\text{sec}$).对某牙医的32个病人,从注射到失去敏感性所需时间在295sec和300sec之间的(近似)概率是多少?

解 ☞ 这里,因为$n \geqslant 30$且试验可无限次重复,故应用无限总体抽样中心极限定理(见13.16节).可以认为$\bar{X}$的分布是近似正态的,参数为$\mu_{\bar{x}}=\mu=290\text{sec}$,$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{25}{\sqrt{32}}=4.419418\text{sec}$.由上述信息,这个问题的近似解是

$$P(295 \leqslant \bar{X} \leqslant 300)$$

由公式(13.15),

$$P(295 \leqslant X \leqslant 300) \approx P\left(\frac{295-290}{4.419418} \leqslant Z \leqslant \frac{300-290}{4.419418}\right)$$

$$\approx P(1.13 \leqslant Z \leqslant 2.26)$$

对$z=1.13$和$z=2.26$查表A.5,

$$P(295 \leqslant \bar{X} \leqslant 300) \approx 0.4881 - 0.3708 = 0.1173$$

用样本和的正态抽样分布计算概率

13.23 例13.6的渔业公司发现,对于它的15只渔船,每船每天的捕捞量服从$\mu=500\text{lb}$和$\sigma=40\text{lb}$的正态分布.如果公司向主要经销商保证每天的捕捞量至少为4200lb,试问:如果某一天只有9只船出海捕鱼,公司不能完成最少保证量的概率是多少?

解 ☞ 问题的本质是,从日捕捞量的正态分布总体抽取$n=9$个日捕捞量($X=x_i$)的一个样本;对9个捕捞量求和$y=\sum_{i=1}^{9} x_i$,然后计算$P(Y<4200)$.因为样本来自正态分布,所以Y的抽样分布也服从正态分布(见13.19节),且$\mu_y=n\mu=9\times500=4500\text{lb}$和$\sigma_y=\sqrt{n}\sigma=\sqrt{9}\times40=120\text{lb}$.由公式(13.19)可知

$$P(Y < a) = P\left(Z < \frac{a-\mu_y}{\sigma_y}\right)$$

因此，

$$P(Y < 4200) = P\left(Z < \frac{4200-4500}{120}\right) = P(Z < -2.50)$$

对 $z=2.50$ 查表 A.5，

$$P(Y < 4200) = 0.5 - 0.4938 = 0.0062$$

用样本和的近似正态抽样分布计算概率

13.24 对例 13.17 的计算机软件公司，上周 1000 个程序员的加班时间的分布有参数 $\mu=5.75$hr和 $\sigma=0.48$hr. 如果抽取 64 个程序员作为一个随机样本(无放回)，上周他们的加班时间的总和至多为 375hr 的概率是多少？

解 ☞ 由于 $n\geqslant30$ 且 $N\geqslant2n$，中心极限定理(见 13.20 节)表明，$Y=\sum_{i=1}^{64}X_i$ 的分布是近似正态的，有参数 $\mu_y=n\mu=64\times5.75=368.0$hr，并且由于 $n>0.05N$，使用公式(13.21)：

$$\sigma_y = \sqrt{n}\sigma\times\sqrt{\frac{N-n}{N-1}} = \sqrt{64}\times0.48\times\sqrt{\frac{1000-64}{1000-1}} = 3.72$$

由上述信息，这个问题的近似解是

$$P(Y\leqslant 375)$$

由公式(13.20)知，

$$P(Y\leqslant375)\approx P\left(Z\leqslant\frac{375-\mu_y}{\sigma_y}\right) = P\left(Z\leqslant\frac{375-368.0}{3.72}\right) = P(Z\leqslant1.88)$$

对 $z=1.88$ 查表 A.5，

$$P(Y\leqslant375)\approx0.5+0.4699 = 0.9699$$

用比率的近似正态抽样分布计算概率

13.25 对例 13.20 的民意测验，调查 100 个选民的一个样本发现，支持 A 的选民比率是 $p=0.49$. 假设在选民总体中支持 A 的比率是 $p=0.51$，则 $P(0.48\leqslant\bar{P}\leqslant0.53)$是多少？

解 ☞ 同例 13.20 一样，假定支持 A 的比率的抽样分布是一个二项分布，且有 $\mu_p=0.51$，$\sigma_p^2=0.002499$和 $\sigma_p=0.05$. 使用连续性修正，$\frac{1}{2n}=\frac{1}{2(100)}=0.005$，由公式(13.32)解得：

$$P(0.48-0.005\leqslant\bar{P}\leqslant0.53+0.005)$$

$$\approx P\left(\frac{0.48-0.005-0.51}{0.05}\leqslant Z\leqslant\frac{0.53+0.005-0.51}{0.05}\right)$$

$$P(0.475\leqslant\bar{P}\leqslant0.535)\approx P(-0.70\leqslant Z\leqslant0.50)$$

对 $z=0.70$ 和 $z=0.50$ 查表 A.5，

$$P(0.48\leqslant\bar{P}\leqslant0.53)\approx0.2580+0.1915 = 0.4495$$

13.26 一个计算机制造商以 10000 的批量购买一种电路晶片. 晶片的 3% 是不合格品时接受此批，但是如果发现来自这批的 200 个晶片的样本中有 5% 或更多的不合格品，则拒绝该批. 他拒绝某一批的概率是多少？

解 ☞ 如果认为这是在 Bernoulli 试验条件下对二项总体的抽样，则不合格品率的抽样分布服从二项分布并有 $\mu_p=p=0.03$，$\sigma_p^2=\frac{pq}{n}=\frac{0.03\times0.97}{200}=0.0001455$ 和 $\sigma_p=\sqrt{\frac{pq}{n}}=\sqrt{0.0001455}=0.012062$. 由于 $np=6$ 和 $nq=194$，可以认为不合格品率的抽样分布近似服从正态分布(见 13.25 节)，并且

$$Z=\frac{\bar{P}-\mu_p}{\sigma_p}=\frac{\bar{P}-0.03}{0.012062}$$

近似服从标准正态分布.因此,使用连续性修正$\frac{1}{2n}=\frac{1}{400}=0.0025$,解[由公式(13.32)]

$$P(\bar{P}\geqslant 0.05-0.0025)=P(\bar{P}\geqslant 0.0475)$$

$$\approx P\left(Z\geqslant\frac{0.05-0.0025-0.03}{0.012062}\right)=P(Z\geqslant 1.45)$$

对 $z=1.45$ 查表 A.5,

$$P(\bar{P}\geqslant 0.05)\approx 0.5-0.4265=0.0735$$

13.27 在某石油公司工作的某地质学家已经确认,在某种半球形岩石结构下有 54% 的机会能发现石油.对于公司在这种岩石下的 100 次钻井,至少有 53% 次发现石油的概率是多少?

解 如果认为这是在 Bernoulli 试验条件下对二项总体的抽样,则成功勘测的比率的抽样分布服从二项分布并有 $\mu_{\bar{p}}=p=0.54$,$\sigma_{\bar{p}}^2=\frac{pq}{n}=\frac{0.54\times 0.46}{100}=0.002484$ 和 $\sigma_{\bar{p}}=\sqrt{\frac{pq}{n}}=\sqrt{0.002484}=0.49840$.由于 $np=54$ 和 $nq=46$,可以认为此比率的抽样分布近似服从正态分布(见 13.25 节),并且

$$Z=\frac{\bar{P}-\mu_{\bar{p}}}{\sigma_{\bar{p}}}=\frac{\bar{P}-0.54}{0.049840}$$

近似服从标准正态分布.因此,使用连续性修正$\frac{1}{2n}=\frac{1}{200}=0.005$,解[从公式(13.32)]

$$P(\bar{P}\geqslant 0.53-0.005)=P(\bar{P}\geqslant 0.525)$$

$$\approx P\left(Z\geqslant\frac{0.53-0.005-0.54}{0.049840}\right)=P(Z\geqslant -0.30)$$

对 $z=0.30$ 查表 A.5,

$$P(\bar{P}\geqslant 0.53)\approx 0.5+0.1179=0.6179$$

补充习题

均值的抽样分布

13.28 一个建筑公司计划修建一条连接两座城市的新公路.为查知期望的交通量,公司从两座城市(根据电话簿)的联合总体无放回抽取各包含 50 个市民的 5 个随机样本;并对每个人询问他们在一个月中使用新公路穿行两座城市的期望次数.以下是 5 个样本(次/月)得到的结果:$\bar{x}_1=6.5$,$\bar{x}_2=8.4$,$\bar{x}_3=7.3$,$\bar{x}_4=7.6$,$\bar{x}_5=6.8$.试问:样本的总平均值是多少?

答案:每月 7.32 次

13.29 对习题 13.4(b)列出的可能随机样本及其均值,计算每一可能均值的被选概率,并且用概率表描述得到的抽样分布.

答案:要求的抽样分布显示在表 13.14 中

表 13.14

样本均值 $\bar{x}$	概率 $f(\bar{x})$
1.5	0.066667
2.0	0.066667
2.5	0.133333
3.0	0.133333
3.5	0.200000
4.0	0.133333
4.5	0.133333
5.0	0.066667
5.5	0.066667
$\sum$	1.000000

13.30 一个数量巨大的甲虫总体(N 基本上无穷)的背部斑点个数彼此不完全相同.可能斑点数是 1,2 或 3 个,并且它们在总体中有相同的比率.从这个斑点个数总体有放回抽取容量 $n=2$ 的所有的可能随机样本,对这些样本确定均值的抽样分布.计算每一可能样本均值的被选概率,并且用概率表描述得到的抽样分布.

答案:要求的抽样分布显示在表 13.15 中

表 13.15

样本均值 $\bar{x}$	概率 $f(\bar{x})$
1.0	0.111111
1.5	0.222222
2.0	0.333333
2.5	0.222222
3.0	0.111111
$\sum$	0.999999

均值的抽样分布的均值,方差和标准差

13.31 表 13.14 中的抽样分布的均值是多少?

答案:3.50

13.32 表 13.14 中的抽样分布的标准差是多少?

答案:1.08

13.33 表 13.15 中的抽样分布的均值是多少?

答案:2.00

13.34 表中 13.15 的抽样分布的标准差是多少?

答案:0.58

用均值的离散型抽样分布计算概率

13.35 对表 13.14 的抽样分布,在一个样本中平均卧室数超过 2.0 间的概率是多少?

答案:0.8667

13.36 对表 13.15 的抽样分布,在一个样本中平均斑点数大于 1.5 但小于 3.0 个的概率是多少?

答案:0.5556

13.37 对习题 13.30 描述的总体以及习题 13.33 和 13.34 的 $\mu_{\bar{x}}$ 和 $\sigma_{\bar{x}}$ 值,$\bar{X}$ 在区间 $\mu_{\bar{x}}\pm1.25\sigma_{\bar{x}}$ 上取值的概率是多少?求解使用:(a)Chebyshev 定理,(b)精确解法(见习题 13.13).

答案:(a)至少 0.3600,(b)0.7778

用均值的正态抽样分布计算概率

13.38 一家出售花生酱的公司声称,每只瓶子装有 18 盎司花生酱.为达到这个标准,公司将灌装机设置为对每只瓶子填充 18.3 盎司.作为质量控制,抽取灌装后的花生酱作为随机样本($n=16$),并称其重量对机器进行例行检查.经过数千次这种检查,公司认为待售花生酱的重量 X 服从正态分布且 $\mu=18.3$oz,$\sigma=0.12$oz.如果样本均值 $\bar{X}=\bar{x}$ 与 $\mu_{\bar{x}}$ 相差超过 3 倍标准差,则对灌装机进行校正.对任一给定的样本,灌装机必须进行校正的概率是多少?

答案:0.0026

13.39 在夏末,雄性地鼠的体重服从正态分布,并且总体均值为 400 克,标准差为 100 克.抽取 50 只地鼠作为一个随机样本.样本均值在 380 克和 420 克之间的概率是多少?

答案:0.8414

13.40 对一个非常庞大的儿童总体,用 Stanford-Binet 智商标准测得的智商得分(IQ 得分)通常服从正态分布,并且总体均值为 100.0,标准差为 16.0.抽取 100 个 IQ 得分作为一个随机样本.样本均值在 102 以上的概率是多少?

答案:0.1056

中心极限定理

13.41 有 500 窝百灵鸟的一个总体,每窝平均产蛋 3.2 个,标准差为 0.8.如果有放回抽取 36 窝作为样本,样本均值大于 3.0 个的概率是多少?

答案:≈0.9332

13.42 如果习题 13.41 描述的百灵鸟总体的抽样($n=36$)是无放回进行,样本均值大于 3 0 个的概率是多少?

答案: ≈0.9406

用样本和的正态抽样分布计算概率

13.43 对例 13.16 和习题 13.23 描述的渔业公司,查知 X_3 的抽样分布服从参数为 $\mu_3=450$lb 和 $\sigma_3=40$lb 的正态分布,如果这对随机样本 $X_1, X_2, \cdots, X_9$ 均成立,那么某一天 9 只船的总捕捞量介于 3850lb 和 3900lb 之间的概率是多少?

答案: 0.0581

用样本和的近似正态抽样分布计算概率

13.44 对于例 13.17 和习题 13.24 描述的计算机软件公司,上周 1000 个程序员的加班时间的分布有参数 $\mu=5.75$hr 和 $\sigma=0.48$hr. 如果抽取 49 个程序员作为一个样本(无放回),上周他们的总加班时间大于等于 280hr 但至多为 283hr 的概率是多少?

答案: ≈0.3428

用比率的近似正态抽样分布计算概率

13.45 如果习题 13.26 的计算机制造商认为从每批抽取一个容量为 200 的样本不可靠,从而将样本容量增加到 400,但是仍然在有 5% 或更多的不合格品时拒绝相应批,则他拒绝某一批的概率是多少?

答案: ≈0.0139

13.46 一个鱼类生物学家经过对若干湖泊的某种鱼总体的大量研究后认为:这种鱼有 44% 是雄性. 如果抽取 40 条这种鱼作为样本,则雄性的比率介于 42% 和 46% 之间的概率是多少?

答案: ≈0.3182

第十四章　总体均值的单样本估计

14.1　估计

推断性统计学是统计科学的一部分，它提供了从样本特征对整个总体（见上册，3.5 节）特征做出推断（一般化）的逻辑和方法．推断性统计学在理论上有 4 个组成部分：概率论（见上册第八章至第十章和本册第十一章和第十二章），抽样理论（见第十三章），估计理论和假设检验理论．本章是推断问题的开始，介绍估计理论在总体均值 μ 的单样本估计中的应用．接着，第十五章处理总体方差 σ^2 和标准差 σ 的估计以及二项总体比例 p 的估计．第十六章介绍单样本假设检验理论．第十七章处理双样本估计和假设检验问题，最后，第十八章讨论这些问题的多样本情形．

估计理论提供了从样本统计量估计未知总体参数的方法．由前知，总体参数是刻画某些总体特征的数值度量，可以从整个测量总体进行计算；而样本统计量（或统计量）是某些测量值样本特征（见上册，3.4 节）的经验性数值度量．不能将经验总体参数和概率分布参数相混淆，后者在理论分布公式中是确定常数．

估计量是指任何一个对总体参数 θ 给出估计值（见 13.8 节）的样本统计量 $\hat{\theta}$．估计值是指从某一样本计算得到的估计量的一个具体数值 $\hat{\theta}^*$．于是，样本均值 $\bar{X}$ 是总体 μ 的一个估计量，对任意给定样本，产生具体估计值 $\bar{X}=\bar{x}$．因为具体的估计值随不同样本变化，所以估计量是随机变量，并有理论分布和经验分布，后者称为抽样分布（见 13.6 节）．

单样本估计问题，是通过来自总体的一个样本对总体参数给出一个数值近似，它的重要性体现在现代生活的各个领域：估计某一人群的选举倾向或电视收视倾向，估计一个野生动物总体的数量，估计总体中某种疾病的流行性，对购买新汽车的总体估计在市区行驶条件下每加仑汽油的行程，等等．

14.2　选择最优估计的标准

数理统计学家使用一些“估计优良性”标准选取总体参数 θ 的最优估计 $\hat{\theta}$，其中有：**无偏性，精密度（简称精度），有效性，一致性**和**充分性**．

13.8 节指出，估计量 $\hat{\theta}$ 是无偏的——具无偏性——如果 $\hat{\theta}$ 的抽样分布的均值 $E(\hat{\theta})$ 等于待估参数：$E(\hat{\theta})=\theta$．同时也指出，绝对距离 $|E(\hat{\theta})-\theta|$ 称为估计量的准确度，当 $|E(\hat{\theta})-\theta|=0$ 时，估计量 $\hat{\theta}$ 称为 θ 的一个完全准确（无偏）的估计量．

无偏性表明，构造这个估计量时不存在系统误差；它既不高估也不低估待估参数．但是如果存在两个或更多的无偏估计量，都不存在系统误差，它们的估计值的离差或变异——包含在估计值中的随机误差——可能不同．估计量的精密度和有效性从不同方面处理这种变异．

13.12 节指出，估计量 $\hat{\theta}$ 的精度是参数 θ 的重复估计值 $\hat{\theta}^*$ 的离差或变异，由 $\hat{\theta}$ 的抽样分布的标准差（标准误）度量．于是，一个估计量称为比另一个更精确，如果它有更小的标准差．

估计量 $\hat{\theta}$ 的效率也与估计的离差有关，是由该估计的**均方误**（MSE）度量，即

$$
\begin{aligned}
(\hat{\theta}\text{ 的 MSE}) &= (\hat{\theta}\text{ 的方差}) + (\hat{\theta}\text{ 的偏}) \\
(\hat{\theta}\text{ 的 MSE}) &= \sigma^2_{\hat{\theta}} + [E(\hat{\theta})-\theta]
\end{aligned}
\tag{14.1}
$$

这表明，MSE 等于估计量的抽样分布的方差加上这个估计量的偏．显然，如果估计量是无偏

的,

$$(\hat{\theta} \text{ 的 MSE}) = \sigma^2_{\hat{\theta}}$$

因此,无偏估计的效率由它的抽样分布的方差来度量,同一参数的两个或更多无偏估计量的效率可以通过比较它们的方差来进行:对两个无偏估计,有较小方差的一个称为**更有效**;多于两个,有最小方差的一个称为**最有效**.

最后两个性质,一致性和充分性,是数学概念,这里我们仅仅给出不严格的定义.估计量 $\hat{\theta}$ 称为一致的——具一致性——如果随着样本容量 n 增加而趋于总体容量 N 时,$\hat{\theta}$ 的估计越来越接近待估参数 θ.对于无偏估计量,有效性,精度和一致性有以下关系:n 愈大,$\sigma^2_{\bar{x}}$ 和 $\sigma_{\bar{x}}$ 愈小,从而估计量变得愈有效且精度也愈高,估计值也越集中(一致)在参数附近.最后,估计量称为充分的,如果它提供包含在样本中有关参数的全部信息.

无偏性和有效性都是一个估计量应该具有的性质.如果参数 θ 存在两个或更多的无偏估计量,则最有效估计——有最小方差——被认为是最优的,并称为**最小方差无偏估计**.于是,尽管存在许多集中趋势估计量可以用于估计总体均值 μ,诸如中位数,众数和中列数(见上册,第6章),但是样本均值 $\bar{X}$ 是 μ 的最小方差无偏估计.类似地,用于估计总体方差 σ^2 的所有离散趋势估计量(见上册,第7章)中,样本方差 S^2 是 σ^2 的最小方差无偏估计.(回顾上册7.7节,样本方差的分母是 $n-1$,使之成为无偏估计.)除了无偏性和有效性,$\bar{X}$ 和 S^2 还都具有一致性和充分性.

样本标准差 S 不是总体标准差 σ 的最小方差无偏估计;而是一个稍微有偏的估计.然而,基于其它"估计优良性"标准,样本标准差总的说来仍不失为 σ 的一个优良估计,你将发现在本书中常常使用.

14.3 均值的估计的标准误 $S_{\bar{x}}$

均值的标准误 $\sigma_{\bar{x}}$(见13.11节)的最优估计是**均值的估计的标准误** $S_{\bar{x}}$,对来自无限或未知容量总体的一个大小为 n 的样本,定义

$$S_{\bar{x}} = \frac{S}{\sqrt{n}} \tag{14.2}$$

如果从有限总体无放回抽样,N 已知,且 $n>0.05N$(见13.10和13.11节):

$$S_{\bar{x}} = \frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \tag{14.3}$$

对给定样本,随机变量 $S_{\bar{x}}$ 有具体值 $s_{\bar{x}}$.由于 S 是 σ 的一个稍微有偏的估计,故 $S_{\bar{x}}$ 也是 $\sigma_{\bar{x}}$ 的一个稍微有偏的估计.然而,基于其它"估计优良性"标准,$S_{\bar{x}}$ 是 $\sigma_{\bar{x}}$ 的优良估计(就像 S 是 σ 的优良估计一样).

14.4 点估计

对于来自一个测量总体的任何随机样本,如果对估计量 $\hat{\theta}$ 算得一个具体数值 $\hat{\theta}^*$ 用以估计总体参数 θ,则数值 $\hat{\theta}^*$ 称为参数 θ 的一个**点估计**.也可以将点估计看做是实轴(见上册,1.20节)上的一个点.

例 14.1 从有限总体抽取一个随机样本,已知:$n=10$,$\sum x_i = 84.1$ 和 $\sum x_i^2 = 708.33$.计算估计量 $\bar{X}$,S^2,S 和 $S_{\bar{x}}$ 的点估计 $\hat{\theta}^*$.

解 对估计量 $\bar{X}$[见上册公式(3.1)和13.5节],

$$X = \frac{\sum X_i}{n}$$

从而点估计是

$$\bar{x} = \frac{84.1}{10} = 8.41$$

对估计量 S^2[见上册公式(7.17)和13.6节],

$$S^2 = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1}$$

从而点估计是

$$s^2 = \frac{708.33 - \frac{(84.1)^2}{10}}{10-1} = \frac{708.33 - 707.281}{9} = 0.116556 \text{ 或 } 0.117$$

对估计量 S[见上册公式(7.24)],

$$S = \sqrt{S^2}$$

从而点估计是

$$s = \sqrt{0.116556} = 0.341403 \text{ 或 } 0.34$$

($\bar{x}$, s^2, s 的近似原则来自上册7.11节.)

对于从无限总体抽样,$S_{\bar{x}}$ 的估计是公式(14.2)

$$S_{\bar{x}} = \frac{S}{\sqrt{n}}$$

从而点估计是

$$s_{\bar{x}} = \frac{0.341403}{\sqrt{10}} = \frac{0.341403}{3.16228} = 0.107961 \text{ 或 } 0.11$$

(标准差点估计舍入到两位有效数字).

14.5　点估计的表示和评价

方差,标准差和均值的标准误的点估计在报告或演讲中常常以小写符号(s^2, s 和 $s_{\bar{x}}$)表示.然而,对样本均值的估计的表示不尽相同,由于专业领域的习惯不同,或者使用 $\bar{X}$ 或者使用 $\bar{x}$,均值表示的不一致,也反映在统计学著作中:一些著作始终使用 $\bar{X}$;另一些著作始终使用 $\bar{x}$;更多的情形,就像这里,在描述性统计学(见上册,第6章和第7章)的讨论中使用 $\bar{x}$,而在介绍随机变量的概念(见上册,10.1节)之后,总是区别变量 $\bar{X}$ 和它的具体取值 $\bar{x}$.这种区别可以使用明确的符号语句,如 $P(\bar{x}_1 < \bar{X} < \bar{x}_2) = 0.95$,表明随机变量 $\bar{X}$ 作为 μ 的估计量,在估计值 $\bar{x}_1$ 和 $\bar{x}_2$ 之间取值的概率是0.95.

反映总体参数的点估计,应该给出尽可能多的附加信息,使得便于评价估计值的"质量":准确度和精度.估计的准确度由包含在估计中的系统误差(偏差)的大小决定.系统误差进入估计过程有以下几种方式:(1)度量方法在某一方向上发生系统失真(见上册,2.13节),(2)抽样设计有缺陷或非随机,导致某种抽样偏差(见上册,3.22节),(3)使用了有偏估计量 $\hat{\theta}$(见13.8和14.2节).估计的精度由随机误差的表现程度——对不同样本,估计值 $\hat{\theta}^*$ 在参数 θ 附近随机变化的大小——决定(见13.12节).

如果样本均值 $\bar{x}$ 是总体均值 μ 的点估计,同时对使用的测量和抽样方法有明确的描述,则可以评价上述系统误差的前两个来源;由于 $\bar{X}$ 是 μ 的无偏估计,故不存在误差的第三个来源.另外,如果已知样本容量 n 和均值的标准误 $\sigma_{\bar{x}}$(通常以 $\bar{x} \pm \sigma_{\bar{x}}$ 给出),并且抽取不同容量的样本对 μ 进行估计,则估计的相对精度可以通过比较 $\sigma_{\bar{x}}$ 的不同值来确定.

通常,总体标准差 σ 是未知的,然而,如果已知样本容量 n,总体容量 N(如果已知的话)以及测量和抽样方法有明确描述,还给出了 $\bar{x}\pm s$ 或 $\bar{x}\pm s_{\bar{x}}$,则统计专业人员可以使用公式(14.2)或(14.3)将 s 变换为 $s_{\bar{x}}$;反之亦然.于是,给定 $\bar{x}\pm s$ 或 $\bar{x}\pm s_{\bar{x}}$,可以:(1)评价所描述的具体随机样本的离差,(2)获得总体标准差 σ 的点估计 s,(3)获得测量均值的标准误 $\sigma_{\bar{x}}$(见13.12节)的精度的点估计 $s_{\bar{x}}$.而且,可以使用上述信息计算总体均值 μ 的**区间估计**——任何统计估计问题的最优完备解.

14.6 点估计和区间估计的关系

对未知总体参数 θ 进行估计可用单个随机样本的点估计或区间估计.由前知,点估计是从样本计算出来的估计量 $\hat{\theta}$ 的一个数值 $\hat{\theta}^*$.尽管单个 $\hat{\theta}^*$ 值不能说明估计的好坏,但也给出一些补充信息:准确度可以通过对度量和抽样方法的描述进行评价;精度可以通过计算估计的标准误(对 $\bar{X}$ 是 $\bar{x}\pm\sigma_{\bar{x}}$)或标准误的估计值(对 $\bar{X}$ 是 $\bar{x}\pm s_{\bar{x}}$)进行评价.

就这一点而言,本章仍然属于描述性统计学的范畴;因为,尽管有点估计及其准确度和精度的一些信息,但是我们仍然未能从样本跳跃到总体.即未能把点估计与待估参数联系起来——给出估计对参数的接近程度或确定在估计值中存在多大的可能误差.为了从样本信息推断总体参数,现在,我们必须进入到推断性统计学领域(见上册,3.5节).

区间估计是一个从样本到总体的推断.以已知的可靠程度,区间估计将总体参数 θ 置于一个实区间 $a\leqslant\theta\leqslant b$ 上,这里 a 和 b 是实数轴上区间的上界和下界.区间的边界值由以下三个因素决定:(1) 样本点估计值 $\hat{\theta}^*$(如 $\bar{x}$),(2) 联系总体参数和样本点估计的样本统计量[如 Z 统计量,$Z=(\bar{X}-\mu)/\sigma_{\bar{x}}$(见13.14节)],(3) 该统计量的抽样分布(如,如果 $\bar{X}$ 服从正态分布,则 Z 统计量的抽样分布是标准正态分布).本章就是讨论如何从单样本计算出这类区间,用以确定总体均值 μ 的位置.这些区间估计通常称为**总体均值的置信区间**.

14.7 导出 $P(\bar{x}_{1-\alpha/2}\leqslant\bar{X}\leqslant\bar{x}_{\alpha/2})=P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$

如果容量为 n 的所有随机样本$(X_1,X_2,\cdots,X_n)$都来自无限大的正态分布总体(均值 μ,方差 σ^2 和标准差 σ),并且对每一样本计算出连续随机变量 $\bar{X}$ 的值 $\bar{x}$,则有

$$P(\bar{x}_{1-\alpha/2}\leqslant\bar{X}\leqslant\bar{x}_{\alpha/2})=P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha \tag{14.4}$$

为理解这一概率表述的推导过程,回顾12.11节知,对连续正态随机变量 X,这个表述的一般形式[公式(12.18)]

$$P(x_{1-\alpha/2}\leqslant X\leqslant x_{\alpha/2})=P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$$

第十二章给出以下几点解释:

(1) 对变量 X,α 表示 X 的正态分布的上尾或下尾的面积,即 x_α 右侧或 $x_{1-\alpha}$ 左侧的面积.

(2) 如果 α 完全落在分布的一侧,并且 X 用 Z 变换 $Z=\dfrac{X-\mu}{\sigma}$ 作了标准化,则 α 表示标准正态分布的上尾或下尾的面积,即 z_α 右侧或 $z_{1-\alpha}$ 左侧的面积.

(3) 如果 α 等分在分布的两侧,用 $\alpha/2$ 表示两个相等面积,则

$$P(x_{1-\alpha/2}\leqslant X\leqslant x_{\alpha/2})=1-\alpha$$

即,X 在 $x_{1-\alpha/2}$ 和 $x_{\alpha/2}$ 之间取值的概率等于分布落在两者之间的面积:$1-\alpha$.

(4) 如果 X 是标准化的,则对标准正态变量 Z 也有

$$P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$$

13.14节指出,如果相同容量 n 的所有可能随机样本$(X_1,X_2,\cdots,X_n)$均来自无限大的正

态分布总本(μ,σ^2,σ),对每一样本计算 $\bar{x}$,则 $\bar{X}$ 的分布(均值的抽样分布)是 $\mu_{\bar{x}}=\mu$ 和 $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}$的正态分布.因此,同上述 X 一样,对 $\bar{X}$ 也成立

$$P(\bar{x}_{1-\alpha/2}\leqslant\bar{X}\leqslant\bar{x}_{\alpha/2})=1-\alpha$$

即 $\bar{X}$ 在 $\bar{x}_{1-\alpha/2}$ 和 $\bar{x}_{\alpha/2}$ 之间取值的概率等于分布落在两者之间的面积:$1-\alpha$.如果 $\bar{X}$ 经 Z 变换作了标准化,则也有

$$P(\bar{x}_{1-\alpha/2}\leqslant\bar{X}\leqslant\bar{x}_{\alpha/2})=P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$$

例 14.2　对标准正态分布,若 $1-\alpha=0.90$,计算 $z_{\alpha/2}$ 和 $-z_{\alpha/2}$.

解　当 $1-\alpha=0.90$ 时,$\alpha=0.10$.由前知[公式(12.16)]:$P(Z>z_{\alpha/2})=\alpha/2$.因此,

$$P(Z>z_{0.10/2})=P(Z>z_{0.05})=0.05$$

于是,查表 A.5 找到 z 值,使得面积最接近于

$$P(0\leqslant Z\leqslant z_{0.05})=0.45$$

得到

$$z_{0.05}=1.645$$

由标准正态分布的对称性,

$$-z_{0.05}=-1.645$$

14.8　导出 $P(\bar{X}-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\mu\leqslant\bar{X}+z_{\alpha/2}\sigma_{\bar{x}})=1-\alpha$

由上节知,如果容量为 n 的所有可能随机样本均来自无限大正态分布总体,并且对每一样本计算出连续随机变量 $\bar{X}$ 的值 $\bar{x}$,则[公式(14.4)]

$$P(\bar{x}_{1-\alpha/2}\leqslant\bar{X}\leqslant\bar{x}_{\alpha/2})=P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$$

对于上述条件,也有

$$P(\bar{X}-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\mu\leqslant\bar{X}+z_{\alpha/2}\sigma_{\bar{x}})=1-\alpha \tag{14.5}$$

这可以通过在不等式中进行算术运算得到(见上册,1.23 节).

首先考虑关系

$$P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$$

代入 $Z=\frac{\bar{X}-\mu}{\sigma_{\bar{x}}}$[公式(13.12)],

$$P\left(-z_{\alpha/2}\leqslant\frac{\bar{X}-\mu}{\sigma_{\bar{x}}}\leqslant z_{\alpha/2}\right)=1-\alpha$$

不等式的三项都乘以 $\sigma_{\bar{x}}$,得到

$$P(-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\bar{X}-\mu\leqslant z_{\alpha/2}\sigma_{\bar{x}})=1-\alpha$$

乘以 -1 改变不等式方向,同时 $\bar{X}-\mu$ 变为 $\mu-\bar{X}$

$$P(z_{\alpha/2}\sigma_{\bar{x}}\geqslant\mu-\bar{X}\geqslant-z_{\alpha/2}\sigma_{\bar{x}})=1-\alpha$$

这等价于

$$P(-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\mu-\bar{X}\leqslant z_{\alpha/2}\sigma_{\bar{x}})=1-\alpha$$

最后,不等式的每一项加上 $\bar{X}$,

$$P(\bar{X}-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\mu\leqslant\bar{X}+z_{\alpha/2}\sigma_{\bar{x}})=1-\alpha$$

上述推导给出了总体均值 μ 的区间估计(见 14.6 节)的概率形式,要求以下条件:容量为 n 的单样本来自无限大且标准差 σ 已知的正态分布总体.这说明,对该总体的任何容量为 n 的**未来**随机样本,$\bar{X}-z_{\alpha/2}\sigma_{\bar{x}}$ 和 $\bar{X}+z_{\alpha/2}\sigma_{\bar{x}}$ 之间的区间包含均值 μ 的概率是 $1-\alpha$.由于 $\bar{X}$ 是随机变量,故这个区间也是随机变量.

例 14.3　随机样本($n=9$, $\bar{x}=14.3$)来自无限大且标准差 σ 已知的正态分布总体.如果 $1-\alpha=0.95$,则 $\alpha=0.05$ 且 $z_{\alpha/2}=z_{0.025}$.查表 A.5 找到面积最接近 0.025 的 z 值是 1.96.因此,在这些条件下,$\sigma_{\bar{x}}=\sigma/\sqrt{n}$.公式(14.5)可以写成

$$P\left(\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\leqslant\mu\leqslant\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha \tag{14.6}$$

在公式(14.6)中代入上述值,得到

$$P\left[14.3-\left(1.96\times\frac{1.5}{\sqrt{9}}\right)\leqslant\mu\leqslant14.3+\left(1.96\times\frac{1.5}{\sqrt{9}}\right)\right]=0.95$$

$$P(13.32\leqslant\mu\leqslant15.28)=0.95$$

试问:上面所做的代入以及最后的概率表述有何差错?

解　概率表述是错误的.未知的总体均值 μ 是一个固定数——不随样本变化.因此,μ 或者在 13.32 到 15.28 的区间上或者不在;即,μ 以非 0 即 1 的未知概率落在区间上.

概率表述

$$P\left(\bar{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\leqslant\mu\leqslant\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

是正确的.它是均值的正态抽样分布和标准正态分布之间关系式的一个导出性质.这说明,对于来自这个总体的随机样本,以 $1-\alpha$ 的概率使 $\bar{X}-z_{\alpha/2}\sigma_{\bar{x}}$ 和 $\bar{X}+z_{\alpha/2}\sigma_{\bar{x}}$ 之间的区间包含 μ.对于给定的具体区间,不能宣称有这个概率;因此,题中错误是在概率表述中用一个具体的样本值 $\bar{x}$ 替换了随机变量 $\bar{X}$.

14.9　总体均值 μ 的置信区间:标准差 σ 已知的正态分布总体

14.6 节指出,一个区间估计以一个已知的确定度将总体参数 θ 置于某指定区间 $a\leqslant\theta\leqslant b$,这里 a 和 b 是实值.同时指出,边界值(a 和 b)取决于:(1)样本点估计 $\hat{\theta}^*$,(2)联系 $\hat{\theta}^*$ 和待估总体参数 θ 的样本统计量,(3)该统计量的抽样分布.在 14.8 节,使用 Z 统计量及其抽样分布(标准正态分布),在已知的可靠程度$(1-\alpha)$下,总体均值 θ 置于[公式(14.5)]

$$P(\bar{X}-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\mu\leqslant\bar{X}+z_{\alpha/2}\sigma_{\bar{x}})=1-\alpha$$

的一个随机区间上.但是,从例 14.3 发现,我们不能使用样本点估计 $\hat{\theta}^*(\bar{x}=14.3)$ 计算概率表述中指定区间的边界值 a 和 b.解决这个难题的途径是计算一个特殊的$(1-\alpha)100\%$**置信区间**,它给出 μ 的正确且完备的区间估计.

为理解置信区间的概念,考虑公式(14.5).它指出,如果容量为 n 的所有样本均来自上述总体(无限大的正态分布总体,σ 已知),并且对每一样本计算随机区间的具体值

$$\bar{X}-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\mu\leqslant\bar{X}+z_{\alpha/2}\sigma_{\bar{x}}$$

则这些确定区间

$$\bar{x}-z_{\alpha/2}\sigma_{\bar{x}}\leqslant\mu\leqslant\bar{x}+z_{\alpha/2}\sigma_{\bar{x}}$$

中有$(1-\alpha)100\%$包含 μ.因此在重复抽样中,我们说有$(1-\alpha)100\%$的**置信度**.上述方法产生的区间包含 μ.基于这个原因,任何这样的区间都称为$(1-\alpha)100\%$置信区间[也写成 $100(1-\alpha)\%$ **置信区间**].于是,如果 $\alpha=0.05$ 即 $1-\alpha=0.95$,则任何这样的区间都将是一个$(0.95)100\%$或 95%置信区间,即有 95% 的信心认为这个区间包含 μ.置信区间估计不是一个概率表述;$(1-\alpha)$只表示这个估计方法的理论成功率.

同点估计一样,置信区间也有估计量和具体估计值.在上述讨论中,随机区间是 μ 的区间估计,由这个估计计算出的任何具体区间即是一个估计值.

例 14.4 在例 14.3 的条件($n=9$, $\bar{x}=14.3$, $\sigma=1.5$)下,如果 $1-\alpha=0.95$,则置信区间估计值是多少?

解 对于 95%置信区间,$z_{\alpha/2}=z_{0.025}$.从例 14.3 知,$z_{0.025}=1.96$.于是,以 95%的置信水平可以说 μ 在区间

$$\bar{x}-1.96\sigma_{\bar{x}} \leqslant \mu \leqslant \bar{x}+1.96\sigma_{\bar{x}}$$

代入 $\bar{x}=14.3$ 和 $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{1.5}{\sqrt{9}}$,

$$14.3-(1.96)\left(\frac{1.5}{\sqrt{9}}\right) \leqslant \mu \leqslant 14.3+(1.96)\left(\frac{1.5}{\sqrt{9}}\right)$$

$$13.32 \leqslant \mu \leqslant 15.28$$

因此,有 95%的置信度认为 μ 在 13.32 和 15.28 之间.

14.10 置信限的表示

以一定的置信水平,我们得到 μ 的一个区间估计 $a \leqslant \theta \leqslant b$,其中 $a=\bar{x}-z_{\alpha/2}\sigma_{\bar{x}}$,$b=\bar{x}+z_{\alpha/2}\sigma_{\bar{x}}$.有$(1-\alpha)100\%$置信度包含 μ 的置信区间的端点值 a 和 b,称为区间的**置信限**.置信区间有以下表示形式:

(1) $\bar{x}-z_{\alpha/2}\sigma_{\bar{x}} \leqslant \mu \leqslant \bar{x}+z_{\alpha/2}\sigma_{\bar{x}}$(同 14.9 节),

(2) $L=\bar{x}-z_{\alpha/2}\sigma_{\bar{x}}$, $U=\bar{x}+z_{\alpha/2}\sigma_{\bar{x}}$,这里 L 表示区间的下端点,U 表示上端点.可能用到的其它符号还有:L(区间的左端点),R(右端点),L_1(下端点或左端点),L_2(上端点或右端点).

(3) $(\bar{x}-z_{\alpha/2}\sigma_{\bar{x}}, \bar{x}+z_{\alpha/2}\sigma_{\bar{x}})$,

(4) $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$.

例 14.5 一篇期刊论文给出总体均值 μ 的置信区间:1.25cm±(1.96×0.078cm).对这个区间,(a)置信限(confidence limits),(b)**信念限**(fiducial limits),(c)**置信系数**,(d)**置信水平**,(e)**置信度**,(f)**临界值**,(g)区间是什么?

解 (a)置信限是

$$L=1.25-(1.96\times 0.078)=1.097\text{cm} \quad U=1.25+(1.96\times 0.078)=1.403\text{cm}$$

(b)一些作者等价地使用信念限和置信限来表示置信区间的端点.

(c)数值 $1-\alpha$ 称为置信系数.因为 $z_{\alpha/2}=1.96$.从例 14.3 知 $\alpha/2=0.025$, $\alpha=0.05$, $1-\alpha=0.95$.因此,置信系数是 0.95.

(d)置信水平是以百分数表示的置信系数.因此,在这里置信水平是 95%.

(e)置信度既指置信水平又指置信系数.因此,在这里置信度是 0.95 或 95%.

(f)临界值是 $\pm z_{\alpha/2}$,即变量 Z 的标准正态分布中面积 $1-\alpha$ 的边界值.因此,在这里临界值 $\pm z_{\alpha/2}$ 是 ± 1.96.

(g)区间是介于置信限之间的实轴上的区间.因此,在这里就是 1.097cm 和 1.403cm 之间的连续区间.

14.11 置信区间的精度

本书到目前为止,已经对术语"精度"给出两个统计学定义(和一个物理学定义,上册 2.15 节).首先(见上册,2.14 节),定义精度为测量的统计性质:重复测量同一物体的接近程度即为该测量的精度.然后(见 13.12 节),定义精度为统计估计量 $\hat{\theta}$ 的一个性质:参数 θ 的重复估计值 $\hat{\theta}'$ 的变异性即为该估计的精度,它由 $\hat{\theta}$ 的抽样分布的标准差(标准误)度量.现在,我们定义精度为置信区间估计的一个性质.

如果容量为 n 的样本重复的取自一个容量为 N 的总体,并且在同一置信水平上对每一样

本计算置信区间,则置信区间估计的精度由这些区间的长度决定:长度越小,区间估计越精确.

对于14.9节的置信区间估计 $\bar{X} \pm z_{\alpha/2}\sigma_{\bar{x}}$,在给定的置信水平和样本容量下,$z_{\alpha/2}$ 和 $\sigma_{\bar{x}}$ 都是常数.因此,所有区间估计值的长度恰是 $2\times(z_{\alpha/2}\sigma_{\bar{x}})$.于是,$2z_{\alpha/2}\sigma_{\bar{x}}$ 决定着该估计的精度;它越小,估计越精确.

尽管 $2z_{\alpha/2}\sigma_{\bar{x}}$ 是估计 $\bar{X} \pm z_{\alpha/2}\sigma_{\bar{x}}$ 的精度的精确度量,但是,许多统计著作将 $z_{\alpha/2}\sigma_{\bar{x}}$ 看做是任何具体的置信区间估计 $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$ 的精度的度量,并且称为 μ **的估计值的误差边缘**或简单地称为**误差边缘**,因为有 $(1-\alpha)100\%$ 的置信度使得抽样误差(见13.11节)的绝对值 $|\bar{x}-\mu|$ 不大于 $z_{\alpha/2}\sigma_{\bar{x}}$.误差边缘用 E 表示:

$$E = z_{\alpha/2}\sigma_{\bar{x}} \tag{14.7}$$

E 越小,估计越精确.一些统计学著作也使用其它表示符号,如 A,B,e,$e_{\alpha/2}$,γ 和 δ.

例 14.6 在例14.4,对于来自无限大正态分布总体($\sigma=1.5$)的一个随机样本($n=9$,$\bar{x}=14.3$),得到如下95%置信区间:

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

$$14.3 \pm (1.96)\left(\frac{1.5}{\sqrt{9}}\right)$$

$$14.3 \pm 0.98$$

现在,对相同数据计算90%和99%置信区间.比较三个置信区间,说明置信水平的增加如何影响置信区间估计的精度.

解 ☞ μ 的95%置信区间估计的精度是[公式(14.7)]$E=0.98$.从例14.2知,对 $1-\alpha=0.90$ 的 $z_{\alpha/2}$ 值是1.645,则90%置信区间是

$$14.3 \pm (1.645)\left(\frac{1.5}{\sqrt{9}}\right)$$

$$14.3 \pm 0.82$$

从而对于90%置信区间,

$$E = 0.82$$

对 $1-\alpha=0.99(\alpha=0.01)$ 求 $z_{\alpha/2}$

$$P(Z > z_{\alpha/2}) = P(Z > z_{0.005}) = 0.005$$

查表A.5,面积最接近0.005的 z 值是2.575.因此,

$$\mu \pm z_{\alpha/2}\sigma_{\bar{x}} = 14.3 \pm (2.575)\left(\frac{1.5}{\sqrt{9}}\right) = 14.3 \pm 1.29$$

从而对于99%置信区间,

$$E = 1.29$$

从以上计算可以看出,如果 $\bar{x}$ 和 $\sigma_{\bar{x}}$ 保持不变,随着置信水平从90%增加到95%再到99%,μ 的置信区间的长度增大,因此区间估计的精度(用 E 度量)降低.实际上,需要做出妥协:要增加区间包含参数的准确度,就不得不接受区间估计的精度的降低.

90%,95%和99%置信区间是最常用的区间,当然,可以使用任意水平的置信区间,目标是要在置信水平和区间估计的精度之间达到最优平衡.

置信区间估计的精度也受样本容量的影响.例如,对 μ 的一个估计 $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$,随着样本容量增加,估计变得越精确,因为估计的误差边缘($E=z_{\alpha/2}\sigma_{\bar{x}}$)减小了.

例 14.7 从例14.3,14.4和14.6的总体($\sigma=1.5$)抽取4个不同容量的随机样本.假设4个样本有相等的样本均值 $\bar{x}=14.3$,通过计算说明:样本容量从 $n=9$ 增加到 $n=18$ 到 $n=36$ 到 $n=900$ 对95%置信区间的影响.

解 ☞ 从例14.6知,对 $n=9$,μ 的95%置信区间是 14.3 ± 0.98.将样本容量从 $n=9$ 增加到 $n=$

18,得到

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

$$14.3 \pm (1.96)\left(\frac{1.5}{\sqrt{18}}\right)$$

$$14.3 \pm 0.69$$

对 $n=36$

$$14.3 \pm (1.96)\left(\frac{1.5}{\sqrt{36}}\right)$$

$$14.3 \pm 0.49$$

对 $n=900$,

$$14.3 \pm (1.96)\left(\frac{1.5}{\sqrt{900}}\right)$$

$$14.3 \pm 0.10$$

于是,随着样本容量从 9 增加到 18 到 36 到 900,μ 的估计的误差边缘从 0.98 减小到 0.69 到 0.49 到 0.10.这表明,如果 σ 已知,对于给定的置信水平,增加样本容量可以使置信区间估计更精确

14.12　已知标准差确定样本容量

一个样本应该多大呢?对任何研究项目这都是一个重要的问题,因为每一次抽样进行测量和分析都要花费财力,时间和其它资源.所以,设置样本容量的目的是:在不对资源造成浪费的情况下,实现研究目标.

如果研究目标是确定未知总体均值 μ 的置信区间估计,则实现这个目标所必需的样本容量是关于所求解的区间估计"质量"的一个函数.这里,区间估计的质量由两部分组成:(1)精度,(2)置信水平.二者都能在公式(14.7):$E=z_{\alpha/2}\sigma_{\bar{x}}$ 中找到,其中,E 是估计的精度的度量,$z_{\alpha/2}$ 是与估计的置信水平有关的临界值.因此,如果 E 和 $z_{\alpha/2}$ 值是事先要求的,并且 σ 已知,在公式(14.7)中用 $\frac{\sigma}{\sqrt{n}}$ 替换 $\sigma_{\bar{x}}$,

$$E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

可由上式解得样本容量 n,能够预先确定满足估计性质的样本容量.确定样本容量的一般求解步骤是:

(1) 对公式两边平方,

$$E^2 = \frac{z_{\alpha/2}^2\sigma^2}{n}$$

(2) 解得 n:

$$n = \frac{z_{\alpha/2}^2\sigma^2}{E^2}$$

或

$$n = \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 \qquad (14.8)$$

这样,如果 σ 已知,并且能够事先决定要求的 E 值(能够容许的估计的误差边缘)和 $z_{\alpha/2}$ 值(要求的置信水平),则由上式可解出达到估计性质所必需的样本容量 n.

例 14.8　设计一个试验,从无限大且标准差 $\sigma=1.5$ 的正态分布总体抽取一个随机样本.试验要求得到

一个95%置信度的置信区间估计,使得 μ 的估计的误差边缘 $E=1.0$.样本容量应该多大?

解 ☞　在公式(14.8)中代入 $\sigma=1.5$, $z_{\alpha/2}=1.96$, $E=1.0$,

$$n=\left(\frac{1.96\times 1.5}{1.0}\right)^2=8.64$$

由于样本容量必须是整数,以上 n 值应该近似到最接近的整数.于是,求解的样本容量是 $n=9$.

14.13　总体均值 μ 的置信区间:来自标准差 σ 已知的任何总体的大样本($n\geqslant 30$)

我们往往不能确知所抽样的有限总体是否服从正态分布.在这种情况下,如果 $n\geqslant 30$ 并且总体容量 N"至少两倍"于样本容量 n(见13.15到13.18节),则由中心极限定理知,X 的抽样分布近似地是一个正态分布.由此,对 $\bar{X}$ 用Z 变换进行标准化[公式(13.12)],

$$Z=\frac{\bar{X}-\mu}{\sigma_{\bar{x}}}$$

则最终 Z 的分布是近似标准正态分布.由于 $\bar{X}$ 和 Z 的这种近似正态性,并且已知 σ,故能够使用14.9节的方法找到 μ 的**近似95%置信区间**.

例14.9　一条州际高速公路上有一段弯曲的下坡路段比较繁忙.假设你是一名公路工程师,正在研究这一繁忙路段是否需要拓宽.你(用雷达仪)测量了经过该路段中点的85辆汽车的行驶速度,得到平均速度 $\bar{x}=66.3$mph.如果从以往的研究中知道 $\sigma=8.3$mph,则经过路段中点的汽车总体的平均速度 μ 的近似95%置信区间是多少?

解 ☞　这里,不知道这些速度的有限总体是否服从正态分布,但是由于 $n\geqslant 30$ 且 N"至少两倍"于85,由中心极限定理,$\bar{X}$ 的抽样分布近似为正态分布.从例14.3知,$z_{\alpha/2}=z_{0.025}=1.96$.这是无放回抽样,但是对于这段繁忙路段来说必然有 $n\leqslant 0.05N$,因此可以使用公式(13.7),

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{8.3\text{mph}}{\sqrt{85}}=0.900261\text{mph}$$

将这些值和 $\bar{x}=66.3$mph 代入 $\bar{x}\pm z_{\alpha/2}\sigma_{\bar{x}}$,得到 μ 的近似95%置信区间:

$$66.3\text{mph}\pm(1.96\times 0.900261\text{mph})$$

$$66.3\text{mph}\pm 1.76\text{mph}$$

14.14　确定总体均值 μ 的置信区间:总体标准差 σ 未知

到本章为止,我们一直接受一个不现实的条件:总体均值 μ 未知时却有已知的总体标准差 σ.这样做是为了用熟知的标准正态分布介绍置信区间.事实上,更一般的情形是 μ 和 σ 均未知,并且有关总体的信息全部来自一个随机样本.幸运的是,此时,仍然有一些统计方法可以用于由样本信息估计 μ.这些方法来自两个领域:**大样本理论**,如果样本比较大($n\geqslant 30$);**小样本理论或精确抽样理论**,如果样本比较小($n<30$).

对于大样本,可以假定:(1)无论总体分布的形式如何,均值的抽样分布近似是正态的(13.15到13.18节),(2)均值的估计的标准误 $S_{\bar{x}}$ 是均值标准误 $\sigma_{\bar{x}}$(见14.3节)的一个合理估计.如果上述假定成立,则比率 $\frac{\bar{X}-\mu}{S_{\bar{x}}}$ 的抽样分布能用 Z 变换 $Z=\frac{\bar{X}-\mu}{\sigma_{\bar{x}}}$(产生标准正态分布)的抽样分布近似,并且 μ 的置信区间能用14.9节使用的方法近似.在14.23和14.24节,我们将进一步讨论大样本方法——何时使用或怎样使用的问题.从14.15到14.22节,我们集中讨论总体服从正态分布,σ 未知且 $n<30$ 时,如何估计 μ.

当 $n<30$ 时,标准正态分布不是比率 $\frac{\bar{X}-\mu}{S_{\bar{x}}}$ 的抽样分布的一个好的近似;而且,随着 n 从30减到更小时,近似变得越来越差.然而,如果能够假设抽样总体服从正态分布,则可以证明:比率 $\frac{\bar{X}-\mu}{S_{\bar{x}}}$ 有一个著名的精确抽样分布,称为 t **分布**或 **student t 分布**.

14.15　t 分布

如果样本容量 n 小于 30,总体标准差 σ 未知,且总体分布能够假设是正态的,则置信区间 $a \leqslant \mu \leqslant b$ 的边界 a 和 b 取决于:(1)样本点估计($\bar{x}$),(2)t 统计量(或 T 统计量),(3)t 统计量的抽样分布,即所谓的 **t 分布**.

t 统计量是随机变量

$$T = \frac{\bar{X} - \mu}{S_{\bar{x}}} \tag{14.9}$$

可以取具体值

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} \tag{14.10}$$

t 分布来自小样本理论,表述为:

如果所有容量为 n 的可能样本均来自均值为 μ,标准差为 σ 的正态分布总体,并且对每一样本计算出随机变量 $T = \frac{\bar{X} - \mu}{S_{\bar{x}}}$ 的值 $t = \frac{\bar{x} - \mu}{s_{\bar{x}}}$,则 t 值有一个连续型概率分布(抽样分布),称为 t 分布.

同前面描述的概率分布一样,t 分布由一个唯一确定的概率函数定义.不做推导,我们给出 t 分布的连续概率密度函数

$$f(t) = c\left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2} \tag{14.11}$$

这里 $t = \frac{\bar{x} - \mu}{s_{\bar{x}}}$

$\nu = n - 1$(符号 ν 是小写希腊字母 nu)

c = 常数,是 ν 的一个函数;使 t 分布的总面积等于 1

可以看出,这个概率函数含有变量 T 的值 t,两个常数 ν 和 c,并且 c 是 ν 的函数.因此,ν 是 t 分布的参数,就像 n 和 p 是二项分布的参数(见 11.9 节)以及 μ 和 σ^2(或 μ 和 σ)是正态分布的参数(见 12.3 节)一样.参数 ν 称为**自由度个数**或简单地称为**自由度**,原因将在 14.17 节讨论.符号 ν 是自由度最常用的符号,不过也可以用 df, $d.f.$, DF, d, r 及其它符号.

同前面的二项和正态分布一样,t 分布也不是单个分布而是一个分布族,每个整数自由度($\nu = n - 1$)即对应一个分布.图 14-1 给出无穷多个 t 分布的其中 4 个,它们对应于 $\nu = 1$, 9, 25, ∞.可以看到,每个分布均是钟形且关于均值 0 对称.并且,每个分布从 $-\infty$ 到 ∞ 连续延

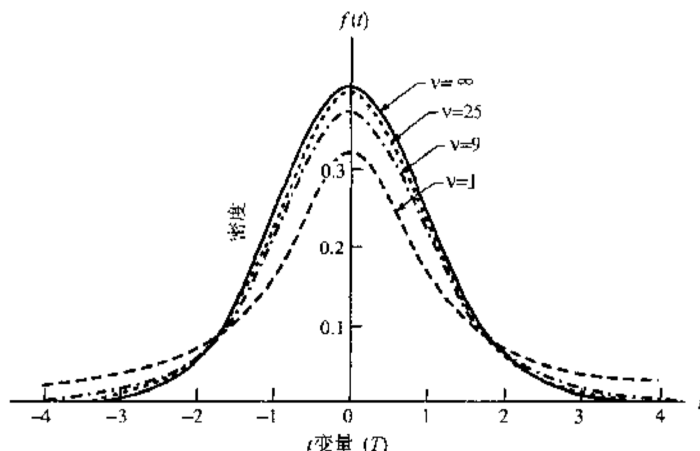


图 14-1

伸,不过这无法在图中显示.

前面指出,数学上推导出 t 分布的概率函数是在正态分布总体抽样的假定下.然而,庆幸的是,已经证实 t 分布是**稳健的**(见上册习题3.14):除非极端偏离正态性,否则 t 分布总是给出有效结果.

14.16 t 分布和标准正态分布的关系

在数学上可以证明,当自由度无限($\nu=\infty$)时,t 分布就是标准正态分布(见12.6节).于是,$\nu=\infty$ 的 t 分布曲线(如图14-1所示)就是标准正态分布曲线.随着自由度从无穷大减小,t 分布曲线保持钟形且关于均值0对称,但是变得比标准正态曲线平坦一些,从而有更多的面积落在尾部.在 $\nu=30$ 之下,t 分布与标准正态分布相差较大.

随着自由度减小,t 分布的离差增加,这反映在 t 分布的标准差上:

$$\sigma=\sqrt{\frac{\nu}{\nu-2}},\ \nu\geqslant 3 \tag{14.12}$$

这样,尽管标准正态分布的标准差总为 $\sigma=1$,但是 $\nu<\infty$ 的 t 分布的标准差总大于1并且随着 ν 减小而增大.随着自由度的减小而标准差增加的趋势可从以下看出:

$$当\ \nu=500\ 时,\quad \sigma=\sqrt{\frac{500}{500-2}}=1.002$$

$$当\ \nu=50\ 时,\quad \sigma=\sqrt{\frac{50}{50-2}}=1.021$$

$$当\ \nu=30\ 时,\quad \sigma=\sqrt{\frac{30}{30-2}}=1.035$$

$$当\ \nu=15\ 时,\quad \sigma=\sqrt{\frac{15}{15-2}}=1.074$$

$$当\ \nu=5\ 时,\quad \sigma=\sqrt{\frac{5}{5-2}}=1.291$$

随着 ν 减小,t 分布比标准正态分布有更大的离差,因为:标准正态分布仅仅反映随机变量 $\bar{X}$ 随样本的变化,但是 t 分布却反映两个随机变量 $\bar{X}$ 和 $S_{\bar{x}}$ 的比例的变化.

14.17 自由度

t 分布的参数 ν 称为自由度.在本书的其余部分,ν 作为不同概率分布的参数还会继续出现,因为,这些分布彼此有联系并且可以相互导出.参数 ν 每次出现均称为自由度的个数,或简单的称为自由度;但是,ν 并不总是用与 t 分布($\nu=n-1$)相同的公式进行计算.因此,必须注意并记住与不同分布相联系的具体的 ν 计算公式.

为全面理解参数 ν 的"自由度"名称的由来,必须理解本书之外的数理统计学的一些概念.不幸的是,对于出现在不同概率分布的参数 ν,根本不存在统一且简单直观的概念解释.然而,对 t 分布恰有一个直观的解释.

与一个统计量相联系的自由度个数可以看做是,在计算该统计量时用到的无限制的、自由变化的数值的个数.例如,如果统计量是样本中 n 个值的和,即 $\sum_{i=1}^{n}X_i$,并且没有任何限制——可以从样本选取任意 n 个值求和,因此,该统计量的自由度为 n.

如果感兴趣的统计量是样本标准差

$$S=\sqrt{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}}$$

在计算中有1个限制：n 个值必须有均值 $\bar{X}=\bar{x}$. 因为首先得计算 $\bar{x}$，所以在计算 $S=s$ 时就减少了一个自由度. 于是，样本标准差(以及样本方差)的自由度是 $n-1$. 可以看到，$n-1$ 出现在 S(和 S^2)统计量的分母中，这也是 S 的分母经常称为自由度的原因，并且 S 的表示式有时也写成

$$S=\sqrt{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{df}}$$

在计算随机变量 T 的具体值时，

$$t=\frac{\bar{x}-\mu}{s_{\bar{x}}}=\frac{\bar{x}-\mu}{s/\sqrt{n}}$$

n 个值也有一个限制：样本均值 $\bar{x}$. 而且它也决定了分子和分母的计算限制. 因此，t 统计量的计算也只有一个限制，从而 t 统计量的自由度为 $\nu=n-1$.

14.18　术语"Student t 分布"

Gosset(1876－1937)是一位英国统计学家，曾经在爱尔兰首都都柏林的 Guinness 啤酒厂工作. 在对啤酒厂进行质量控制的研究中，Gosset 意识到需要进行小样本统计推断. 发现 t 分布是他的众多贡献之一. 当时啤酒厂有规定，禁止雇员将研究成果公开发表，于是，Gosset 在1908年的论文中，偷偷地以笔名"Student"发表了 t 分布的发现. 正是由于这个原因，t 分布也称为 **Student t 分布**.

14.19　t 分布的临界值

对于标准正态分布，12.10 节[公式(12.9)]$P(Z>z_\alpha)=\alpha$ 指出，在 z_α 右边标准正态曲线下方的面积(概率)是 α. 因此，值 z_α 是标准正态分布的一个**临界值**[见习题 14.5(f)]；也称为**临界边界值**，因为它是左边 $1-\alpha$ 面积和右边 α 面积之间的边界点. 12.10 节[公式(12.11)]也指出，由于标准正态分布的对称性，$P(Z<-z_\alpha)=\alpha$. t 分布有类似的临界值. 即，对自由度为 ν 的 t 分布，

$$P(T>t_{\alpha,\nu})=\alpha \tag{14.13}$$

这里，T 是随机变量 $T=\frac{\bar{X}-\mu}{S_{\bar{x}}}$，$t_{\alpha,\nu}$ 表示自由度为 ν 的 t 分布中右边 α 和左边 $1-\alpha$ 之间的边界. 面积和单侧临界值之间的这种关系如图 14-2(a)所示. 与标准正态分布一样，t 分布也对称，因此

$$P(T<-t_{\alpha,\nu})=\alpha \tag{14.14}$$

这里 $-t_{\alpha,\nu}$ 表示左边 α 和右边 $1-\alpha$ 之间的边界. 对于自由度为 ν 的 t 分布，面积和单侧临界值之间的这种关系如图 14-2(b)所示.

对于标准正态分布，前面也处理过双侧概率，即 α 在两个尾部等分，可以表示为[公式(12.18)]

$$P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$$

这表明：随机变量 Z 在 $-z_{\alpha/2}$ 和 $z_{\alpha/2}$ 之间取值的概率等于 $1-\alpha$. 这样，$-z_{\alpha/2}$ 和 $z_{\alpha/2}$ 值就是双侧临界值. 类似地，对自由度为 ν 的 t 分布，

$$P(-t_{\alpha/2,\nu}\leqslant T\leqslant t_{\alpha/2,\nu})=1-\alpha \tag{14.15}$$

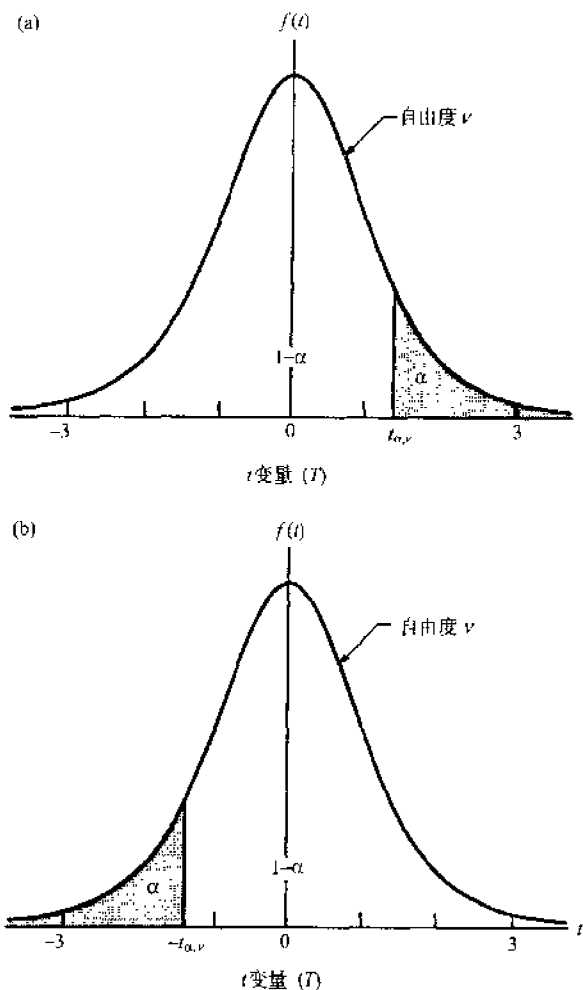


图 14-2

这表明:随机变量 $T=\frac{\overline{X}-\mu}{S_{\bar{x}}}$在 $-t_{\alpha/2,\nu}$和 $t_{\alpha/2,\nu}$之间取值的概率等于 $1-\alpha$. 同样, $-t_{\alpha/2,\nu}$和 $t_{\alpha/2,\nu}$值是自由度为ν的t分布的双侧临界值. t分布的面积和双侧临界值之间的这种关系如图 14-3 所示.

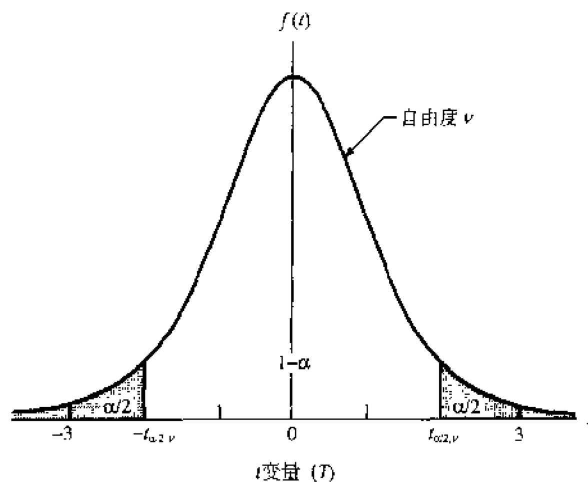


图 14-3

14.20　表 A.6: t 分布的临界值

对一些挑选出的 t 分布($\nu=1,\cdots,30,40,60,120,\infty$),附录表 A.6(**$t$ 分布的临界值**)给出 $t_{\alpha,\nu}$ 和 $t_{\alpha/2,\nu}$ 的正临界值.前面章节所列的表只给出一个概率分布的某定义域内的面积(概率),比较而言,表 A.6 有新格式:每行表示一个不同的 t 分布,由 ν 的具体值决定;每行左边是自由度 ν 值,以后是分布右(正)尾指定面积的临界值;这个右尾面积就是每列顶端的符号 t_α 中的 α 值.

表 14.1 是表 A.6 的一部分,给出 5 个 t 分布:$\nu=21, 22, 23, 24, 25$.临界值的第一列是每一分布的右尾面积 $\alpha=0.10$ 的临界值 $t_{0.10}$.临界值的第二列是右尾面积 $\alpha=0.05$ 的临界值 $t_{0.05}$,等等.因为 t 分布关于均值 0 对称,

$$P(T<-t_{\alpha,\nu})=P(T>t_{\alpha,\nu})=\alpha \tag{14.16}$$

表 14.1

ν	t_α					
	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.0005}$
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	(2.060)	2.485	2.787	3.725

这表明:对于自由度为 ν 的 t 分布,$-t_{\alpha,\nu}$ 左侧面积等于 $t_{\alpha,\nu}$ 右侧面积,且均为 α.虽然表 A.6 只给出正临界值,但是同样可找到负临界值,因为除符号外 $-t_{\alpha,\nu}$ 和 $t_{\alpha,\nu}$ 相等.

例 14.10　查表 A.6 的部分表 14.1,找到 $t_{\alpha,\nu}=t_{0.025,25}$ 和 $-t_{\alpha,\nu}=t_{0.025,25}$.

解 ☞　对于自由度 $\nu=25$ 的 t 分布,使得右侧面积为 $\alpha=0.025$(见图 14-4)的值 $t_{0.025,25}=2.060$(表 14.1 中划圈值).由公式(14.16):$-t_{0.025,25}=-2.060$.

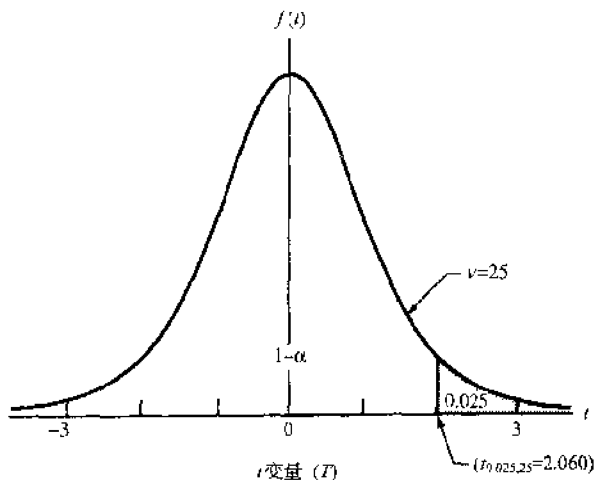


图 14-4

表 A.6 的设计不仅可以找到单侧临界值,而且也可找到双侧临界值.此时,α 被 t 分布的两个尾部等分,每个尾部都有 $\alpha/2$.由上知,表 A.6 的每一行给出一个不同的 t 分布(由 ν 决定),每一列给出一个在右尾截下 α 面积的正临界值 $t_{\alpha,\nu}$,在表 A.6 中,一列中的每个值都是不同 ν 值的一个 $t_{\alpha/2}$.

例 14.11　查表 A.6 的部分表 14.2,求 $t_{\alpha/2,\nu}=t_{0.05/2,25}$ 和 $-t_{\alpha/2,\nu}=-t_{0.05/2,25}$.

解 问题是要求在自由度 $\nu=25$ 的 t 分布中找到一个临界值,使得在该值右侧的面积为 $\alpha/2=0.05/2=0.025$. 在表 14.2 中找到 $\nu=25$ 的行,沿该行找到 $t_{0.05/2}$ 的列,得到值 2.060(划圈). 这个值将分布分成右侧等于 $\alpha/2$(即 0.025)的一个面积和左侧等于 $1-\alpha/2$ 的一个面积. 这与例 14.10 所求 t 值 $t_{\alpha,\nu}=t_{0.025,25}$ 相同,因为两个临界值都在分布的右尾截下 0.025 的面积. 由于 t 分布关于均值 0 对称,$t_{\alpha/2,\nu}$ 和 $t_{\alpha/2,\nu}$ 除符号外是相等的,因此 $-t_{0.025,25}=-2.060$. 对给定的 t 分布,t_α 和 $t_{\alpha/2}$ 之间的关系如图 14-5 所示.

表 14.2

ν	$t_{0.20/2}$	$t_{0.10/2}$	$t_{0.05/2}$	$t_{0.02/2}$	$t_{0.01/2}$	$t_{0.001/2}$
	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.0005}$
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	(2.060)	2.485	2.787	3.725

(表头:$t_{\alpha/2}$ 上层,t_α 下层)

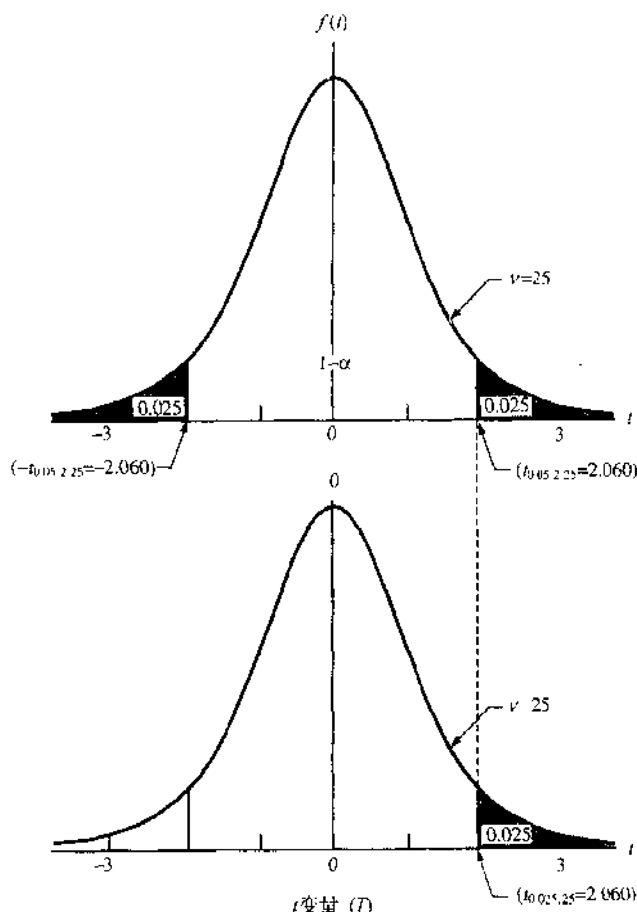


图 14-5

14.21 总体均值 μ 的置信区间:来自标准差 σ 未知的正态总体的小样本($n<30$)

现在,讨论正态分布的总体均值 μ 的置信区间,假设标准差未知且样本容量小于 30. 对这样一个小样本,不能假定:(1)均值的抽样分布近似正态,(2)均值的估计标准误 $S_{\bar{x}}$ 是均值标准误 $\sigma_{\bar{x}}$ 的一个合理估计. 因此,我们不能将均值的抽样分布变换为标准正态分布,也不能用

$\pm z_{\alpha/2}$作为临界值.在上述条件下,改用$\pm t_{\alpha/2,\nu}$作为临界值,总体均值μ的精确区间估计,即$(1-\alpha)100\%$置信区间,由下式计算

$$\bar{x}-t_{\alpha/2,\nu}s_{\bar{x}}\leqslant\mu\leqslant\bar{x}+t_{\alpha/2,\nu}s_{\bar{x}} \tag{14.17}$$

为理解上式,回顾14.15节,对容量为n的任意样本:如果容量为n的所有可能样本均来自均值为μ,标准差为σ的正态分布总体,并且对每一样本计算出随机变量$T=\frac{\bar{X}-\mu}{S_{\bar{x}}}$的值$t=\frac{\bar{x}-\mu}{s_{\bar{x}}}$,则$t$值有一个称为$t$分布的连续概率分布.而且,这个$t$分布有参数(自由度)$\nu=n-1$.

14.19节表明,以下概率表述[公式(14.15)]对来自正态分布总体的任意容量的样本均严格成立:

$$P(-t_{\alpha/2,\nu}\leqslant T\leqslant t_{\alpha/2,\nu})=1-\alpha$$

用$\frac{\bar{X}-\mu}{S_{\bar{x}}}$替换$T$[公式(14.19)],

$$P\left(-t_{\alpha/2,\nu}\leqslant\frac{\bar{X}-\mu}{S_{\bar{x}}}\leqslant t_{\alpha/2,\nu}\right)=1-\alpha$$

对不等式执行与14.18节相同的算术运算,导出区间估计的以下概率形式

$$P(\bar{X}-t_{\alpha/2,\nu}S_{\bar{x}}\leqslant\mu\leqslant\bar{X}+t_{\alpha/2,\nu}S_{\bar{x}})=1-\alpha \tag{14.18}$$

于是可以认为,如果所有容量为n(小或大)的样本均来自μ和σ未知的正态分布总体,并且对每一样本计算出随机区间

$$\bar{X}-t_{\alpha/2,\nu}S_{\bar{x}}\leqslant\mu\leqslant\bar{X}+t_{\alpha/2,\nu}S_{\bar{x}}$$

的具体值.那么,具体区间

$$\bar{x}-t_{\alpha/2,\nu}s_{\bar{x}}\leqslant\mu\leqslant\bar{x}+t_{\alpha/2,\nu}s_{\bar{x}}$$

以$(1-\alpha)100\%$包含总体均值μ.因此,有$(1-\alpha)100\%$的置信度使得任何由此确定的区间包含μ,并且任何这样的具体区间均是$(1-\alpha)100\%$置信区间.同14.10节一样,置信区间有以下表示形式:

(1) $\bar{x}-t_{\alpha/2,\nu}s_{\bar{x}}\leqslant\mu\leqslant\bar{x}+t_{\alpha/2,\nu}s_{\bar{x}}$(如上),

(2) $L=\bar{x}-t_{\alpha/2,\nu}s_{\bar{x}}$, $U=\bar{x}+t_{\alpha/2,\nu}s_{\bar{x}}$,

(3) $(\bar{x}-t_{\alpha/2,\nu}s_{\bar{x}}, \bar{x}+t_{\alpha/2,\nu}s_{\bar{x}})$,

(4) $\bar{x}\pm t_{\alpha/2,\nu}s_{\bar{x}}$.

例14.12　一个容量为$n=16$的随机样本来自μ和σ未知的正态分布总体.如果样本有均值$\bar{x}=27.9$和标准差$s=3.23$,则μ的95%置信区间是多少?

解　由于这是一个小样本,我们使用t分布的置信区间

$$\bar{x}\pm t_{\alpha/2,\nu}s_{\bar{x}}$$

为此需要得到

$$s_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{3.23}{\sqrt{16}}=0.807500$$

和表A.6中所要求的置信水平下的$t_{\alpha/2,\nu}$值.对95%的置信水平,$1-\alpha=0.95$,$\alpha=0.05$和$\alpha/2=0.05/2$.由于$\nu=n-1=16-1=15$,得到

$$t_{\alpha/2,\nu}=t_{0.05/2,15}=2.131$$

从而95%置信区间是

$$27.9 \pm (2.131 \times 0.08075)$$

$$27.9 \pm 1.72$$

14.22　确定样本容量:来自标准差 σ 未知的正态总体的小样本

对于总体标准差 σ 未知的情形,14.11 节定义 μ 的估计的误差边缘为[公式(14.7)]:$E=z_{\alpha/2}\sigma_{\bar{x}}$.由这个定义,14.12 节表明得到要求 E 值所必需的样本容量可以由公式(14.8):$n=\left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$ 确定.

如果总体是正态的,σ 未知且 $n<30$,得到要求 E 值所必需的样本容量是一个较为复杂的问题.现在,合适的置信区间是 $\bar{x}\pm t_{\alpha/2}s_{\bar{x}}$(见 14.21 节),$\mu$ 的估计量的误差边缘定义如下:

$$E=t_{\alpha/2,\nu}s_{\bar{x}} \tag{14.19}$$

$t_{\alpha/2,\nu}$ 和 $s_{\bar{x}}$ 都依赖于具体样本,因此,用这个公式计算 E 的可能途径只能是:取定样本后,计算出置信区间.基于这两个量都是随样本而不同,要得到某一要求 E 值,不能用

$$n=\left(\frac{t_{\alpha/2,\nu}s}{E}\right)^2$$

确定所必需的样本容量.建议使用的方法是,将以前研究过的样本标准差 s 代入近似公式

$$n\approx\left(\frac{z_{\alpha/2}s}{E}\right)^2 \tag{14.20}$$

从而得到一个粗略的 n 值.

一般,这个近似解将低估 n.因为所用的 $z_{\alpha/2}$ 等价于 $t_{\alpha/2,\infty}$(见 14.16 节),对给定的 α,$z_{\alpha/2}$ 是可能的最小 $t_{\alpha/2}$ 值.

例 14.13　对例 14.12 的 95%置信区间,样本为 $\bar{x}=27.9$ 和 $s=3.23$,在重复研究中,n 多大时一个样本才能使得 μ 的估计的误差边缘为 $E=1.25$?

解　为找到能使 $E=1.25$ 所需的 n 值,将 $s=3.23$,$z_{\alpha/2}=z_{0.05/2}=1.96$ 和 $E=1.25$ 代入公式(14.20),

$$n\approx\left(\frac{1.96\times3.23}{1.25}\right)^2=25.7$$

因此,为使 $E=1.25$,样本容量应至少为 26.

例 14.14　如果在例 14.13 的研究中,样本容量 $n=26$,则对于 95%置信区间,要求的 $E=1.25$ 能达到吗?如果不能,多大的样本容量将达到这个水平?

解　如果 n 为 26,则

$$s_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{3.23}{\sqrt{26}}=0.633455$$

$$t_{\alpha/2,\nu}=t_{0.05/2,25}=2.060$$

于是

$$E=t_{\alpha/2,\nu}s_{\bar{x}}=2.060\times0.633455=1.30492 \text{ 或 } 1.30$$

它接近但稍大于 $E=1.25$.为找到合适的样本容量,需要对较大的样本计算 E.

对 $n=27$,

$$E=t_{0.05/2,26}s_{\bar{x}}=2.056\times\frac{3.23}{\sqrt{27}}=1.27804 \text{ 或 } 1.28$$

对 $n=28$,

$$E=t_{0.05/2,27}s_{\bar{x}}=2.052\times\frac{3.23}{\sqrt{28}}=1.25257 \text{ 或 } 1.25$$

对 $n=29$,

$$E=t_{0.05/2,28}s_{\bar{x}}=2.048\times\frac{3.23}{\sqrt{29}}=1.22838 \text{ 或 } 1.23$$

由于为达到 $E=1.25$ 必需抽取容量为 28 的样本,故计算出样本容量为 26 的近似公式将容量低估了 2. 因为存在这种低估,为达到期望的 E 值,建议将近似的 n 值增加 10% 到 15%.

14.23　总体均值 μ 的置信区间:来自标准差 σ 未知的正态总体的大样本($n\geqslant30$)

为获得总体均值的置信区间,当总体服从正态分布且标准差未知时,对从中抽取的一个任意容量($n\geqslant30$ 或 $n<30$)的样本,t 分布方法在理论上给出一个严格的精确解. 即严格的置信区间由 $\bar{x}\pm t_{\alpha/2,\nu}s_{\bar{x}}$ 给出.

然而,对于大样本($n\geqslant30$)的区间估计,统计学著作没有统一的最好方法. 若使用 t 分布的精确解问题在于容量 $n\geqslant30$ 的样本,精确的 $t_{\alpha/2,\nu}$ 值或许无法获得. 大部分统计学著作收录的表,如表 A.6 那样,仅给出 $\nu=1,\cdots,30,40,60,120$ 和 ∞ 的 $t_{\alpha/2}$ 值. 尽管有更完整的 t 分布表,并且可通过计算机编程得到精确值 $t_{\alpha/2,\nu}$,但在 $n\geqslant30$ 时常常只要近似 t 值. 使用近似值的两个建议是:(1)用表中的最接近的较小自由度对应的 t 值,(2)用表中的 t 值插值.

例 14.15　使用以上两个建议,查表 A.6 计算 $t_{0.05/2,50}$ 的近似解.

解　根据建议(1),对于表 A.6 中最接近的较小自由度,t 值是

$$t_{0.05/2,50}\approx t_{0.05/2,40}=2.021$$

根据建议(2),使用**线性插值**方法. 首先构造平行标尺,自由度作为一个标尺,$t_{0.05/2,\nu}$ 值作为另一个标尺,如图 14-6 所示. 自由度 $\nu=50$ 在自由度标尺的位置(标有叉号)是从 $\nu=40$ 到 $\nu=60$ 距离的$(10/20=\frac{1}{2})$. 因此,$t_{0.05/2,50}$ 在 $t_{0.05/2,\nu}$ 标尺上(也标有叉号)也是 2.021 和 2.000 之间距离的$\frac{1}{2}$. 于是,

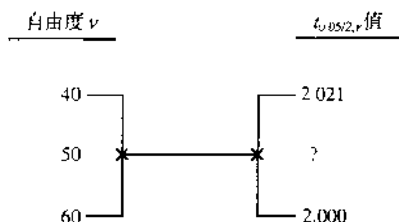


图 14-6

$$t_{0.05/2,50}\approx\frac{1}{2}(2.021-2.000)+2.000=2.0105 \text{ 或 } 2.011$$

对 $n\geqslant30$,一个更简单的方法是使用近似 t 值,一些统计学著作给出以下规则:

如果 σ 未知且抽样总体服从正态分布,那么当 $n\geqslant30$(或在某些著作,$n>30$)时,总可使用 $\bar{x}\pm z_{\alpha/2}S_{\bar{x}}$. 这个解法也得到一个近似$(1-\alpha)100\%$置信区间,但是推荐使用的作者认为,在这些条件下用 z 值来近似精确的 t 值有充足的理由,因为:(1)如果总体是正态的,则均值的抽样分布也总是正态的,(2)对 $n\geqslant30$,均值的估计标准误 $S_{\bar{x}}$ 是均值标准误 $\sigma_{\bar{x}}$ 的一个合理估计,(3)当 $n=30$ 以上时,对相同的 α,值 $z_{\alpha/2}$ 和 $t_{\alpha/2,\nu}$ 之间几乎没有差别.

例 14.16　$n=41$ 的一个随机样本来自 σ 未知的正态分布总体. 样本的均值是 $\bar{x}=12.3$,标准差是 $s=0.90$. 由以上信息,分别用 $\bar{x}\pm z_{0.05/2}s_{\bar{x}}$ 和 $\bar{x}\pm t_{0.05/2,\nu}s_{\bar{x}}$ 估计总体均值 μ.

解　由上述规则知,在本题的条件下,可以使用近似的 95% 置信区间 $\bar{x}\pm z_{0.05/2}s_{\bar{x}}$ 估计 μ. 这里

$$\bar{x}=12.3$$

$$z_{0.05/2}=1.96$$

$$s_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{0.90}{\sqrt{41}}=0.140556$$

从而近似 95% 置信区间是

$$12.3\pm(1.96\times0.140556)$$

$$12.3\pm0.275$$

由上述讨论也知,在理论上,本题的精确 95% 置信区间是 $\bar{x} \pm t_{0.05/2,\nu}s_{\bar{x}}$. 查附录表 A.6 得到精确值 $t_{0.05/2,10}=2.021$. 因此,精确 95% 置信区间是

$$12.3 \pm (2.021 \times 0.140556)$$

$$12.3 \pm 0.284$$

注: 尽管两个区间非常接近,但精确区间更宽,从而精度稍低. 因此,精确区间称为 μ 的比较保守的估计.

14.24　总体均值 μ 的置信区间:来自标准差 σ 未知的非正态总体的大样本($n \geqslant 30$)

精确 t 分布方法要求对正态分布总体抽样. 然而,在大样本($n \geqslant 30$)情形,即使总体的标准差 σ 未知,也没有必要使用具有正态性假定的 t 分布. 因为对于大样本:(1) 中心极限定理指出,均值的抽样分布是近似正态的,(2) 均值的估计标准误 $S_{\bar{x}}$ 是均值标准误 $\sigma_{\bar{x}}$ 的一个合理估计. 因此,对于来自非正态分布总体的大样本,近似置信区间仍然可以是 $\bar{x} \pm z_{\alpha/2}s_{\bar{x}}$.

例 14.17　一家便民连锁店昼夜 24 小时服务,正在考察一家新店的选址. 措施之一是:选取连续的 60 个晚上,计算从晚上 10 点到早晨 5 点的交通流量(从任一方向经过选址处的汽车数). $n=60$ 的样本有 $\bar{x}=238.2$ 和 $s=31.32$ 因为周末交通繁忙,交通流量的总体有偏斜而是非正态. σ 未知,试计算总体均值 μ 的 95% 置信区间.

解 ☞　样本容量为 60,使用 $\bar{x} \pm z_{\alpha/2}s_{\bar{x}}$ 可以计算出近似 95% 置信区间. 这里 $\bar{x}=238.2$, $z_{0.05/2}=1.96$,并且由于 $n \leqslant 0.05N$,

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{31.32}{\sqrt{60}} = 4.04339$$

因此,近似 95% 置信区间是

$$238.2 \pm (1.96 \times 4.04339)$$

$$238.2 \pm 7.93$$

14.25　总体均值 μ 的置信区间:来自非正态总体的小样本($n<30$)

当从非正态总体抽取一个小样本($n<30$)时,估计问题不存在令人满意的标准解法. 表 14.3 是本章内容的一个概括. 表中的适当的置信区间(精确或近似)作为 σ 是否已知,总体是否为正态分布和样本容量的一个函数. 注意,对于表中标有问号的两行——都是关于非正态分布总体的小样本,Z 分布和 t 分布方法都不合适.

一个可行但极少使用的方法与 Chebyshev 定理(见上册,7.15 节)有关.

对任意 $k \geqslant 1$ 以及一个数据集 $x_1, x_2, \cdots, x_n$(或 $x_1, x_2, \cdots, x_N$),测量值落在它们均值的 k 倍标准差范围内的比例至少是 $1-\frac{1}{k^2}$.

对于总体均值的抽样分布,Chebyshev 定理的概率表述为

$$P(|\bar{X}-\mu| \leqslant k\sigma_{\bar{x}}) \geqslant 1-\frac{1}{k^2}$$

这表明,总体均值 μ 与任意样本均值 $\bar{x}$ 的距离在 $k\sigma_{\bar{x}}$ 之内的概率至少是 $1-\frac{1}{k^2}$. 因此,对于来自任意总体的任意容量的样本,如果 σ 已知,则至少有 $1-\frac{1}{k^2}$ 的置信度可以认为 μ 落在近似区间*

$$\bar{x} \pm k\sigma_{\bar{x}}$$

* 此处原文有误.——译者注

表 14.3

总体	样本容量	置信区间	精确或近似	定义的章节
已知 σ				
正态	$n\geqslant30$	$\bar{x}\pm z_{\alpha/2}\sigma_{\bar{x}}$	精确	14.9
正态	$n<30$	$\bar{x}\pm z_{\alpha/2}\sigma_{\bar{x}}$	精确	14.9
非正态	$n\geqslant30$	$\bar{x}\pm z_{\alpha/2}\sigma_{\bar{x}}$	近似	14.13
非正态	$n<30$	?	?	?
未知 σ				
正态	$n\geqslant30$	$\bar{x}\pm t_{\alpha/2,\nu}s_{\bar{x}}$	对精确的 t 值,精确	14.23
		or		
		$\bar{x}\pm z_{\alpha/2}s_{\bar{x}}$	近似	14.23
正态	$n<30$	$\bar{x}\pm t_{\alpha/2,\nu}s_{\bar{x}}$	精确	14.21
非正态	$n\geqslant30$	$\bar{x}\pm z_{\alpha/2}s_{\bar{x}}$	近似	14.24
非正态	$n<30$	?	?	?

之内;如果 σ 未知,则至少有 $1-\frac{1}{k^2}$ 的置信度可以认为 μ 落在近似区间

$$\bar{x}\pm ks_{\bar{x}}$$

之内.

例 14.18　一种新的心脏手术正在一家医院推广.对于已完成的 20 例这种手术,平均住院期为 $\bar{x}=14.3$ 天($s=2.84$ 天).因为手术复杂,住院期天数的总体不服从正态分布,而是有些正偏.σ 未知,使用 Chebyshev 定理计算总体均值 μ 的至少 90%近似置信区间.

解 ☞　首先令 $1-\frac{1}{k^2}$ 等于 0.90,解得 k:

$$1-\frac{1}{k^2}=0.90$$

$$\frac{1}{k^2}=0.10$$

$$0.10k^2=1$$

$$k=\sqrt{10}=3.16228$$

由于 σ 值未知,将 k 值代入 $\bar{x}\pm ks_{\bar{x}}$,这里

$$\bar{x}=14.3\text{ 天}$$

$$s_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{2.84\text{ 天}}{\sqrt{20}}=0.635043\text{ 天}$$

于是,μ 的至少 90%的近似置信区间是

$$14.3\text{ 天}\pm(3.16228\times0.635043\text{ 天})$$

$$14.3\text{ 天}\pm2.01\text{ 天}$$

如果可以假设该总体是正态的,即能够使用 t 分布方法,则可以得到有更高精度的精确 90%置信区间:

$$14.3\text{ 天}\pm[(t_{0.10/2,19}=1.729)\times(0.635043\text{ 天})]$$

$$14.3\text{ 天}\pm1.10\text{ 天}$$

习题解答

点估计

14.1　一家制衣厂收到 150 个制服纽扣,纽扣为红色,直径 10cm.为确认纽扣直径是否为 10cm,无放回抽出 9 个纽扣作为随机样本,测得直径分别为(精确至毫米):9.9,10.9,

10.0,10.1,9.9,10.1,10.0,10.0,9.9.根据此样本,计算 $\bar{X}$,S 和 $S_{\bar{x}}$ 的点估计.

解

$$\bar{X} = \frac{\sum X_i}{n}$$

$$\bar{x} = \frac{89.9}{9} = 9.98889\text{cm} \text{ 或 } 9.99\text{cm}$$

$$S = \sqrt{\frac{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1}}$$

$$s = \sqrt{\frac{898.05 - \frac{(89.9)^2}{9}}{9-1}} = \sqrt{\frac{898.05 - 898.0011}{8}}$$

$$= 0.0781825\text{cm} \text{ 或 } 0.078\text{cm}$$

因为是从有限总体无放回抽样,且 $n>0.05N$,由公式(14.3),

$$S_{\bar{x}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

$$S_{\bar{x}} = \frac{0.0781825}{\sqrt{9}}\sqrt{\frac{150-9}{150-1}}$$

$$= 0.0260608 \times 0.972784 = 0.0253515\text{cm} \text{ 或 } 0.025\text{cm}$$

总体均值的置信区间:标准差已知的正态总体

14.2 对标准正态分布,当 $1-\alpha=0.99$ 时,求 $z_{\alpha/2}$ 和 $-z_{\alpha/2}$.

解 由于 $1-\alpha=0.99$,$\alpha=0.01$,因此,

$$P(Z > z_{\alpha/2}) = P(Z > z_{0.01/2}) = P(Z > z_{0.005}) = 0.005$$

对于面积最接近

$$P(0 \leqslant Z \leqslant z_{0.005}) = 0.495$$

的对应 z 值,查表 A.6 得

$$z_{0.005} = 2.575$$

从而

$$-z_{0.005} = -2.575$$

14.3 一个随机样本来自无限大的正态分布总体,总体标准差 $\sigma=1.5$.抽样前希望有 98% 的置信度使得 μ 的估计的误差边缘为[公式(14.7)]$E=0.85$.试问应该抽取多大的一个样本?

解 由题知 $1-\alpha=0.98$,$\alpha=0.02$,$\alpha/2=0.01$.因此,查表 A.6 找到 z 值,使得面积最接近 0.49,$z_{\alpha/2}=z_{0.01}=2.33$.将该值与 $E=0.85$,$\sigma=1.5$ 一起代入公式(14.8),

$$n = \left(\frac{2.33 \times 1.5}{0.85}\right)^2 = 16.91$$

向上取整得 $n=17$.于是,欲使估计的误差边缘以 98% 的置信度为 $E=0.85$,所需样本容量为 $n=17$.

14.4 一篇期刊论文认为一个无限大正态分布总体的均值 μ 的 95% 置信区间为 1.25cm ± (1.96×0.078cm).样本容量 $n=25$,总体标准差 σ 未知.若要使估计量的误差边缘以 99% 的置信度为 $E=0.1$cm,则样本容量应该多大?

解 已知 $n=25$,首先需要从 95% 置信区间计算 σ 值:

$$1.25\text{cm} \pm (1.96 \times 0.078\text{cm})$$

$$z_{\alpha/2}\sigma_{\bar{x}} = 1.96 \times 0.078\text{cm}$$

由于 $n \leqslant 0.05N$，可由公式(13.7)计算 σ：

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 0.078\text{cm}$$

$$\sigma = 0.078\text{cm}\sqrt{25} = 0.39\text{cm}$$

其次，把 $\sigma = 0.39\text{cm}$，$E = 0.1\text{cm}$ 和 99%置信水平的临界值 $z_{\alpha/2} = z_{0.005} = 2.575$ 代入公式(14.8)，

$$n = \left(\frac{2.575 \times 0.39}{0.1}\right)^2 = 100.85$$

取整得 $n = 101$. 因此，若要以 99%置信度使估计的误差边缘为 0.1cm，样本容量应为 $n = 101$.

14.5 假设你为某种子公司开发出一种快速生长的洋葱新品种. 现在，你想确定该种洋葱从播种到成熟(可从外观上判断球茎发育，顶端弯曲等)所需的平均时间 μ(天数). 假定从初步的研究知道，平均时间服从 $\sigma = 8.3$ 天的正态分布. 为了获得一个精度为 $E = 2$ 天的 μ 的 95%置信区间，你需要抽取多少个成熟期测量值作为样本？

解 ☞　将 $\sigma = 8.3$ 天，$z_{0.025} = 1.96$，$E = 2$ 天代入等式(14.8)，

$$n = \left(\frac{1.96 \times 8.3}{2}\right)^2 = 66.16$$

向上取整得 $n = 67$.

14.6 在习题 14.5 的洋葱研究中，你抽取了 67 个成熟期作为样本，且样本均值 $\bar{x} = 71.2$ 天. 95%置信区间是多少?

解 ☞　已知 $n = 67$ 和 $\sigma = 8.3$ 天. 因此，

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{8.3\text{天}}{\sqrt{67}} = 1.01401\text{ 天}$$

将该值，$z_{0.025} = 1.96$ 和 $\bar{x} = 71.2$ 天代入 $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$，得到 95%置信区间：

$$71.2\text{ 天} \pm (1.96 \times 1.01401\text{ 天})$$

$$71.2\text{ 天} \pm 1.99\text{ 天}$$

14.7 社会学家正在研究居住在乡村的 10—12 岁儿童的看电视习惯. 从 $N = 39200$ 个乡村儿童中无放回抽取 $n = 50$ 个儿童作为随机样本，得到每周看电视的平均时间为 $\bar{x} = 12.5$ 小时. 假设儿童每周看电视时间的总体服从正态分布，且已知 $\sigma = 2.2$ 小时，则儿童每周平均看电视时间 μ 的 96%置信区间是多少?

解 ☞　由于 $1 - \alpha = 0.96$，则 $\alpha = 0.04$，$\alpha/2 = 0.02$. 查表 A.5，面积最接近 0.48 对应的 z 值为 $z_{\alpha/2} = z_{0.02} = 2.05$. 尽管是从有限总体无放回抽样，但是由于 $n \leqslant 0.05N$，所以可使用公式(13.7)计算 $\sigma_{\bar{x}}$：

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{2.2}{\sqrt{50}} = 0.311127\text{ 小时}$$

将该值与 $z_{0.02} = 2.05$，$\bar{x} = 12.5\text{hr}$ 一起代入 $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$，得到 96%置信区间：

$$12.5\text{ 小时} \pm (2.05 \times 0.311127\text{ 小时})$$

$$12.5\text{ 小时} \pm 0.64\text{ 小时}$$

14.8 对习题 14.7 的研究，若要使 96%置信区间的 E 值为上述解的一半，则样本容量 n 应取多大?

解 ☞　上题 $E = 0.64\text{hr}$，因此须找到 n 使得 $E = 0.32\text{hr}$. 代入 $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$，由公式(14.8)解得，

$$n=\left(\frac{2.05\times 2.2}{0.32}\right)^2=198.63$$

向上取整得 $n=199$. 于是,要使 E 值变为原来的一半,样本应扩大到原来的 4 倍.

14.9 将一种白天用的感冒药溶入盛有 6 盎司水的一个容器,其中感冒药所列成分之一是 200mg guaifenesin. 作为质量控制,制造商抽取 n 个容器作为样本,并逐一分析 guaifenesin 和其它成分的含量. 然后对每种成分计算总体均值 μ 的一个 99.9% 置信区间. 如果制造商已知 guaifenesin 含量服从正态分布且 $\sigma=0.11$mg. 如果制造商对 guaifenesin 的 99.9% 置信区间期望有 $E=0.05$mg, 则样本容量 n 必须为多大?

解 由于 $1-\alpha=0.999$, 则 $\alpha=0.001$. $\alpha/2=0.0005$. 所以,对最接近 0.4995 的面积,查表 A.5 找到 z 值. 因为与 0.4995 对应有 6 个 z 值,取它们的平均:

$$z_{\alpha/2}=z_{0.0005}=\frac{3.27+3.28+3.29+3.30+3.31+3.32}{6}=3.295$$

将该值与 $E=0.05$mg, $\sigma=0.11$mg 一起代入公式(14.8):

$$n=\left(\frac{3.295\times 0.11}{0.05}\right)^2=52.55$$

向上取整得 $n=53$.

总体均值的置信区间:来自标准差已知的任何分布的大样本

14.10 假设你是例 14.9 所述的公路工程师,对经过公路某处的 85 辆汽车测得平均速度是 $\bar{x}=66.3$mph. 从以前研究知道 $\sigma=8.3$mph. 你决定用一个更小的样本重新进行速度研究,并期望得到 95% 置信区间的 E 值为例 14.9 的 $E=1.76$mph 值的 2 倍. 必须抽取多大的样本?

解 因为假设 $\bar{X}$ 近似服从正态分布,并且可以使用 $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}$(见例 14.9),所以利用近似公式(14.8):

$$n\approx\left(\frac{z_{\alpha/2}\sigma}{E}\right)^2 \tag{14.21}$$

求解该问题 代入 $z_{\alpha/2}=z_{0.025}=1.96$, $\sigma=8.3$mph 和 $E=(2\times1.76=3.52\text{mph})$,

$$n\approx\left(\frac{1.96\times 8.3}{3.52}\right)^2=21.36$$

向上取整得到答案:需要一个容量为 22 的样本. 然而,由于这里使用了中心极限定理,故最小样本容量 n 必须至少为 30.

14.11 假设你在一家航空公司工作,并且正在研究乘客离机之后取回行李所需时间. 你已知道,所需时间的总体不服从正态分布,而是有些正偏,$\sigma=6.4$min. 抽取 45 个乘客作为一个随机样本,得到平均时间 $\bar{x}=26.7$min. 总体的平均时间 μ 的 99% 置信区间是多少?

解 因为 $n\geqslant30$ 且 N"至少两倍"于 45*,由中心极限定理,$\bar{X}$ 近似服从正态分布,并且由于 $n\leqslant0.05N$,故可以使用公式(13.7)计算 $\sigma_{\bar{x}}$:

$$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{6.4}{\sqrt{45}}=0.954056\text{min}$$

将该值与 $z_{\alpha/2}=z_{0.005}=2.575$, $\bar{x}=26.7$min 一起代入 $\bar{x}\pm z_{\alpha/2}\sigma_{\bar{x}}$,即可得到 μ 的近似 99% 置信区间:

$$26.7\text{min}\pm(2.575\times0.954056\text{min})$$

$$26.7\text{min}\pm2.46\text{min}$$

---

* 此处原文有误. ——译者注

t 分布

14.12　查附录表 A.6 求:(a)$t_{\alpha,\nu}=t_{0.10,6}$和(b)$-t_{\alpha,\nu}=-t_{0.10,15}$.

解 ☞　(a)$t_{\alpha,\nu}=t_{0.10,6}=1.440$,(b)$-t_{\alpha,\nu}=-t_{0.10,15}=-2.602$

14.13　查附录表 A.6 求:(a)$-t_{\alpha/2,\nu}=-t_{0.001/2,60}$和(b)$t_{\alpha,\nu}=t_{0.0005,40}$.

解 ☞　(a)$-t_{\alpha/2,\nu}=-t_{0.001/2,60}=-3.460$,(b)$t_{\alpha,\nu}=t_{0.0005,40}=3.551$

14.14　用表 A.6,对以下递增的自由度:5,15,30,60,120,∞分别计算 $t_{\alpha,\nu}=t_{0.025,\nu}$.

解 ☞

$$t_{0.025,5}=2.571$$
$$t_{0.025,15}=2.131$$
$$t_{0.025,30}=2.042$$
$$t_{0.025,60}=2.000$$
$$t_{0.025,120}=1.980$$
$$t_{0.025,\infty}=1.960$$

可以看出,对给定的 α 值(这里 $\alpha=0.025$),随着自由度 ν 增加,$t_{\alpha,\nu}$值减小.这是因为,当 ν 增加时,分布曲线下方的面积从尾部向均值(见图 14-1)移动,因此从均值到边界 $t_{\alpha,\nu}$有更小的距离.14.16 节指出,当 $\nu=\infty$时 t 分布变成标准正态分布,这可以从表 A.6 的 $\nu=\infty$行看出.本题中,$t_{0.025,\infty}=1.960$,而从表 A.5 可以看出,在右侧截下面积 $\alpha=0.025$ 的值恰好是 $z_{0.025}=1.96$.

14.15　用表 A.6,对以下递减的 α 值:0.025,0.01,0.005,0.0005 分别计算 $t_{\alpha,15}$.

$$t_{0.025,15}=2.131$$
$$t_{0.01,15}=2.602$$
$$t_{0.005,15}=2.947$$
$$t_{0.0005,15}=4.073$$

可以看出,对给定的 t 分布,$t_{\alpha,\nu}$随着α 减小而增加.

14.16　查表 A.6 求 $t_{1-\alpha,\nu}=t_{0.99,18}$.

解 ☞　对 t 分布有:

$$t_{1-\alpha,\nu}=-t_{\alpha,\nu} \tag{14.22}$$

表明:对自由度为 ν 的 t 分布和给定的 α 值,在右侧有 $1-\alpha$ 面积的 t 值 $t_{1-\alpha,\nu}$与在左侧有 α 面积的 t 值 $-t_{\alpha,\nu}$相等.查表 A.6 知

$$t_{\alpha,\nu}=t_{0.01,18}=2.552$$

因此.

$$-t_{\alpha,\nu}=-t_{0.01,18}=-2.552$$

从而

$$t_{1-\alpha,\nu}=t_{0.99,18}=-2.552$$

总体均值的置信区间:来自标准差未知的正态总体小样本

14.17　一个容量为 $n=16$ 的随机样本取自 μ 和σ 均未知的正态分布总体.如果样本有均值 $\bar{x}=27.9$,标准差 $\sigma=3.23$,μ 的 99%置信区间是多少?

解 ☞　因为是一个小样本,t 分布的置信区间是合适的(见 14.21 节):$\bar{x}\pm t_{\alpha/2,\nu}s_{\bar{x}}$.对 99%置信水平,$1-\alpha=0.99$,$\alpha=0.01$,$\alpha/2=0.01/2$.由于 $\nu=n-1=15$,查表 A.6 找到:$t_{\alpha/2,\nu}=t_{0.01/2,15}=2.947$.从正态分布总体抽样,有[公式(14.2)]$s_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{3.23}{\sqrt{16}}=0.80750$,从而 99%置信区间是

$$27.9 \pm (2.947 \times 0.8075)$$

$$27.9 \pm 2.38$$

14.18 一位专门从事人类进化研究的人类学家在非洲某地区发现7具成年的直立行走猿人的骨骸，这类猿人骨骸以前从未在该地区发现过. 人类学家测量了它们的头盖骨容量(头骨的大脑区域)，以立方厘米 cm^3 为单位, 得到以下结果：925，892，900，875，910，906 和 899. 对抽样总体的均值 μ，计算95%置信区间.

解 头盖骨容量通常服从正态分布，因此这里也如此假定. 尽管这是从有限总体无放回抽样，但是由于不知道与样本相关的总体的容量，所以，使用公式(14.2)、$\frac{s}{\sqrt{n}}$ 计算95%置信区间，即有 $\bar{x} \pm t_{\alpha/2,\nu}\left(\frac{s}{\sqrt{n}}\right)$. 这里

$$\bar{x} = \frac{\sum x}{n} = \frac{6307}{7} = 901.0\text{cm}^3$$

$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{1\,444}{6}} = 15.5134\text{cm}^3$$

$$\sqrt{n} = \sqrt{7} = 2.64575$$

$$t_{\alpha/2,\nu} = t_{0.05/2,6} = 2.447$$

因此，μ 的99%置信区间是

$$901.0\text{cm}^3 \pm \left[2.447 \times \left(\frac{15.5134\text{cm}^3}{2.64575}\right)\right]$$

$$901.0\text{cm}^3 \pm 14.35\text{cm}^3$$

14.19 对习题14.18的人类学家的研究中，为了使95%置信区间的 $E = 10.0\text{cm}^3$，他必须找到多少块头骨?

解 在公式(14.20)中代入习题14.18计算出的标准差 $s = 15.5134\text{cm}^3$，$z_{\alpha/2} = 1.96$ 和 $E = 10.0\text{cm}^3$，

$$n \approx \left(\frac{1.96 \times 15.5134}{10.0}\right)^2 = 9.25$$

向上取整为 $n = 10$，并增加15%(见例14.14)得到 $n = 12$. 因此，为使得 $E = 10.0\text{cm}^3$，人类学家必须另外找到至少5块头骨.

14.20 一个飞机制造公司收到一批轴承，批量为400个，每个轴承的直径规定为1.000cm. 负责人随机抽取了25个作为样本，如果样本的 $\bar{x}$ 是1.000cm，并且从样本计算出的99.9%置信区间有0.001cm或更小的 E 值，则接受该批. 样本数据在表14.4中给出，该批会被接受吗?

解 由上册公式(6.9)，

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i} = \frac{25.001}{25} = 1.00004\text{cm}$$

假定这直径总体服从正态分布，则99.9%置信区间是 $\bar{x} \pm t_{0.001/2,24} s_{\bar{x}}$. 查表A.6，$t_{0.001/2,24} = 3.745$. 由上册公式(7.35)，

$$s = \sqrt{\frac{\sum f_i x_i^2 - n\bar{x}^2}{n-1}} = \sqrt{\frac{(25.002025)^2 - 25(1.00004)^2}{24}} = 0.001020\text{cm}$$

因为这是来自有限总体的无放回抽样，而且 $n > 0.05N$，可用公式(14.3)计算 $s_{\bar{x}}$：

$$s_{\bar{x}} = \frac{S}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{0.001020}{\sqrt{25}}\sqrt{\frac{400-25}{400-1}} = 0.000198\text{cm}$$

因此,99.9%置信区间是

$$1.00004\text{cm} \pm (3.745 \times 0.000198\text{cm})$$

$$1.00004\text{cm} \pm 0.000742\text{cm}$$

这样,$E = 0.000742$cm 达到了"$E = 0.001$cm 或更小"的要求.两个标准都已达到,接受这一批.

表 14.4

直径(cm)x_i	频数 f_i	$f_i x_i$(cm)	x_i^2(cm²)	$f_i x_i^2$(cm²)
0.998	2	1.996	0.996004	1.992008
0.999	4	3.996	0.998001	3.992004
1.000	12	12.000	1.000000	12.000000
1.001	5	5.005	1.002001	5.010005
1.002	2	2.004	1.004004	2.008008
$\sum$	25	25.001cm		25.002025cm²

总体均值的置信区间:来自标准差未知的正态总体大样本

14.21 一个心理学家正在研究出生时服用药物的婴儿的运动肌能力的发育.她采用的一项指标是婴儿能够坐立时的年龄.观察 50 个这样的婴儿后,她发现 $\bar{x} = 129.4$ 天,$s = 15.26$天.σ 未知,分别使用 $\bar{x} \pm z_{0.01/2} s_{\bar{x}}$ 和 $\bar{x} \pm t_{0.01/2,\nu} s_{\bar{x}}$ 计算 μ 的 99%置信区间.

解 通常,运动肌能力的发育时间服从正态分布,因此这里也如此假定.由于 $n \geq 30$,σ 未知,从 14.23 节知,本题的近似解是 $\bar{x} \pm z_{\alpha/2} s_{\bar{x}}$.这里

$$\bar{x} = 129.4 \text{ 天}$$

$$z_{0.01/2} = 2.575$$

而且,由于在理论上这是一个可无限次重复的试验,

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{15.26 \text{ 天}}{\sqrt{50}} = 2.15809 \text{ 天}$$

于是,近似 99%置信区间是

$$129.4 \text{ 天} \pm (2.575 \times 2.15809 \text{ 天})$$

$$129.4 \text{ 天} \pm 5.56 \text{ 天}$$

为计算精确的 99%置信区间 $\bar{x} \pm t_{0.01/2,\nu} s_{\bar{x}}$,需要得到精确的 $t_{0.01/2,49}$ 值.这个值不能从表 A.6 查到.但是,可以借鉴例 14.15 求精确 t 值所建议的两种方法.

根据建议(1),表 A.6 中对最接近的较小自由度的 t 值是 $t_{0.01/2,40} = 2.704$.因此,99%置信区间是

$$129.4 \text{ 天} \pm (2.704 \times 2.15809 \text{ 天})$$

$$129.4 \text{ 天} \pm 5.84 \text{ 天}$$

根据建议(2),使用线性插值,构造平行标尺(见图 14-7):一个是自由度,另一个是 $t_{0.01/2,\nu}$值.由于自由度 $\nu = 49$ 是从 $\nu = 40$ 到 $\nu = 60$ 距离的 9/20(标有叉号),所以 $t_{0.01/2,49}$也是 $t_{0.01/2,\nu}$标尺上2.704和 2.660 之间距离的 9/20(标有叉号).于是

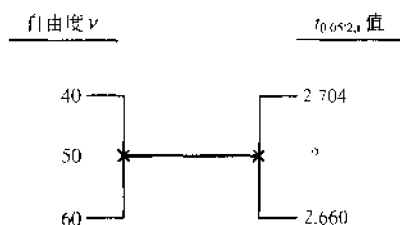


图 14-7

$$t_{0.01/2,49} \approx 2.704 - \left[\frac{9}{20} \times (2.704 - 2.660 = 0.044)\right] = 2.684$$

因此,99%近似置信区间是

$$129.4 \text{ 天} \pm (2.684 \times 2.15809 \text{ 天})$$

$$129.4 \text{ 天} \pm 5.79 \text{ 天}$$

14.22 一个 $n = 70$ 的随机样本取自 μ 和 σ 未知的正态分布总体.样本均值是 22,标准差是

8. 用 $\bar{x} \pm z_{\alpha/2} s_{\bar{x}}$ 计算 μ 的近似 95% 置信区间.

解 ☞

$$\bar{x} = 22$$

$$z_{0.05/2} = 1.96$$

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{8}{\sqrt{70}} = 0.956183$$

近似 95% 置信区间是

$$22 \pm (1.96 \times 0.956183)$$

$$22 \pm 1.87$$

总体均值的置信区间:标准差未知的非正态分布

14.23 一个健康保险公司想了解某城镇的 50000 个居民的平均寿命. 年龄总体的均值和标准差都未知. 公司假定该总体非正态分布,而是有些正偏. 抽取 100 个居民作为一个随机样本,得到平均年龄 $\bar{x} = 40$ 年,标准差 $s = 15$ 年. 计算总体的近似的 95% 置信区间.

解 ☞ 样本容量为 100,我们使用 $\bar{x} \pm z_{\alpha/2} s_{\bar{x}}$ 计算近似 95% 置信区间. 这里

$$\bar{x} = 40\text{ 年}$$

$$z_{0.05/2} = 1.96$$

因为 $n \leqslant 0.05N$,

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{15\text{ 年}}{\sqrt{100}} = 1.5\text{ 年}$$

因此,近似 95% 置信区间是

$$40\text{ 年} \pm (1.96 \times 1.5\text{ 年})$$

$$40\text{ 年} \pm 2.94\text{ 年}$$

14.24 海洋生物学家想了解一种濒危海豚的身长. 对于这种海豚,总体均值和总体标准差均未知. 生物学家知道,在过去的十年中这种海豚没有繁殖,因此他假设海豚身长的分布是负偏的,与分布是正态的情形相比有更多的成年个体. 他只能测量到 10 个海豚,得到平均身长 $\bar{x} = 2.667$ 米,标准差 $s = 0.820$ 米. 使用 Chebyshev 定理(见 14.25 节)计算总体均值 μ 的至少 95% 近似置信区间.

解 ☞ 第一步是求 k:

$$1 - \frac{1}{k^2} = 0.95$$

$$\frac{1}{k^2} = 0.05$$

$$k^2 = 20$$

$$k = 4.47214$$

然后,将 k 值代入 $\bar{x} \pm ks_{\bar{x}}$,这里

$$\bar{x} = 2.667\text{ 米}$$

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{0.820\text{ 米}}{\sqrt{10}} = 0.259307\text{ 米}$$

则 μ 的至少 95% 近似置信区间是

$$2.667\text{ 米} \pm (4.47214 \times 0.259307\text{ 米})$$

$$2.667\text{ 米} \pm 1.1597\text{ 米}$$

补充习题

点估计

14.25　一个随机样本($n=12$)取自无限总体,并且计算出以下和:$\sum_{i=1}^{12}x_i=64$ 和 $\sum_{i=1}^{12}x_i^2=362$.估计量 $\bar{X}$,S^2、S 和 $S_{\bar{x}}$ 的点估计分别是多少?

答案:$\bar{x}=5.3,s^2=1.88,s=1.4,s_{\bar{x}}=0.4$

置信区间:标准差已知的正态分布

14.26　随机样本($n=12,\bar{x}=5.3$)取自无限大正态分布总体,且标准差 $\sigma=1.4$,对 $1-\alpha=0.95$,置信区间估计是什么?

答案:$4.51\leqslant\mu\leqslant6.09$

14.27　对习题14.6的快速生长的洋葱品种,包含67个成熟期的样本有均值 $\bar{x}=71.2$ 天,如果假设成熟期服从 $\sigma=8.3$ 天的正态分布,总体均值 μ 的97%置信区间是什么?

答案:71.2天±2.20天

14.28　如果在习题14.27的洋葱研究中,数据来自含有82个成熟期的一个样本,总体均值 μ 的97%置信区间是什么?

答案:71.2天±1.99天

14.29　随机样本($n=25,\bar{x}=80$)取自无限大正态分布总体,总体标准差已知为 $\sigma=10$.对95%和99%置信区间,计算估计值的误差边缘.

答案:对95%置信区间,$E=3.92$;对99%置信区间,$E=5.15$

14.30　两个随机样本($n=20$ 和 $n=40$)取自同一个无限大正态分布总体,总体标准差已知为 $\sigma=10$.两个样本有相同均值 $\bar{x}=100$.两个样本的99%置信区间分别是多少?

答案:对 $n=20$,100 ± 5.76;对 $n=40$,100 ± 4.07

置信区间:来自标准差已知的任意分布的大样本

14.31　假设你是一个开发商,想了解社区内200000座房屋的平均成本.你随机抽取60座房屋作为样本,得到均值 $\bar{x}=\$65000$.从以前研究知道房屋成本的测量值总体是正偏的,并有标准差 $\sigma=\$7000$.总体均值 μ 的近似95%置信区间是什么?

答案:\$65000±\$1771.2

14.32　在习题14.31的房屋研究中,如果60座房屋的样本来自一个800座房屋的总体,同样有 $\bar{x}=\$65000$ 和 $\sigma=\$7000$,总体均值 μ 的近似95%置信区间是什么?

答案:\$65000±\$1704.6

t 分布

14.33　查表A.6求 $t_{\alpha/2,\nu}=t_{0.01/2,8}$.

答案:3.355

14.34　查表A.6求 $t_{-\alpha/2,\nu}=t_{-0.02/2,8}$.

答案:-1.330

置信区间:来自标准差未知的正态分布小样本

14.35　容量 $n=20$ 的一个随机样本取自 μ 和 σ 未知的正态分布总体.如果样本有均值 $\bar{x}=8.36$ 和标准差 $\sigma=0.92$,则 μ 的95%置信区间是什么?

答案:8.36±0.431

14.36　对习题14.35的95%置信区间,使用 $s=0.92$,求近似的 n 使得均值 μ 的估计的误差边缘达到 $E=0.25$.

答案:$n\approx52.02$ 或至少53

14.37　容量 $n=16$ 的一个随机样本取自 μ 和 σ 未知的正态分布总体.如果样本有均值 $\bar{x}=27.9$,标准差 $\sigma=3.23$,则 μ 的90%置信区间是什么?

答案:27.9±1.42

14.38　16个女志愿者在6个月内每次餐前服用一种节食剂.对这个16人的样本计算出6个月内减掉的体重(镑数),并用 t 分布方法计算出平均减掉体重 μ 的98%置信区间:8.3lb±0.162625lb.样本的 $\bar{x}$ 和 s 是多少?

答案：$\bar{x}=8.3$lb，$s=0.25$lb

14.39　如果重复习题 14.38 的节食剂研究，为使得 98% 置信区间有 $E=0.10$lb，样本应该多大？

答案：至少 40

置信区间：来自标准差未知的正态分布的大样本

14.40　$n=9$ 的随机样本取自 μ 和 σ 均未知的正态分布总体. 样本均值是 $\bar{x}=600$、标准差是 $s=120$. 使用 t 分布和插值法计算近似 95% 置信区间.

答案：600 ± 25.0

14.41　对习题 14.40 的随机样本与总体，用标准正态分布计算近似 95% 置信区间.

答案：600 ± 24.7

置信区间：标准差未知的非正态分布

14.42　假设你想知道林地中一大块橡树总体的平均高度 μ. 已知橡树高度总体有负偏而并非正态，即与正态分布相比，有更多的高大成年橡树. 你测量了 60 棵橡树（$n\leqslant0.05N$）的高度并得到：$\bar{x}=20.5$ 米，$s=7.81$ 米. μ 的近似 95% 置信区间是什么？

答案：$20.5\text{m}\pm1.98\text{m}$

14.43　假设你一直在测量一个沙漠鼠总体的体重. 去年这个总体数量增加了两倍多，而且因为这种迅速增加，存在较多的体重较轻的幼鼠个体（正偏），你认为这个体重总体并不服从正态分布. 总体的 μ 和 σ 均未知. 为估计 μ，你测量了 20 个沙漠鼠的一个样本，并得到：$\bar{x}=17.01$g，$s=5.671$g. 使用 Chebyshev 定理计算 μ 的至少 90% 近似置信区间.

答案：$17.01\text{g}\pm4.010\text{g}$

第十五章　总体方差、标准差及比率的单样本估计

15.1　方差、标准差及比率的最优估计

我们在十四章介绍了由单样本估计总体均值 μ 的统计问题的各种置信区间解. 这些方法是基于两个基本分布：正态分布和 t 分布. 本章我们用单样本来确定其它三个参数：正态分布总体的方差 σ^2 和标准差 σ 及二项分布总体比率 p 的置信区间估计. 如十四章那样，我们将上述每个总体参数 θ 置于区间 $a \leqslant \theta \leqslant b$ 中，其中 a 和 b 分别为实数轴上区间的上下界. 这些边界值是由以下条件决定的(见第 14.6 节)：(1) 样本的点估计值 θ^*，(2)使点估计值与总体参数有关的样本统计量，(3) 该统计量的抽样分布.

在 14.2 节，我们回顾了选择总体参数 θ 的最优估计量 $\hat{\theta}$ 的标准：无偏性，相容性，有效性，一致性及充分性，还指出了总体方差 σ^2 的最优估计量为样本方差 S^2，总体标准差 σ 的最优估计量为样本标准差 S(尽管其是有偏的)，又知任意特定的样本估计值 $S^2 = s^2$ 是 σ^2 的一个点估计值. 任意特定的样本估计值 $S = s$ 是 σ 的一个点估计值，现在我们说二项分布参数 p 的最优估计量是样本比率 $\overline{P}$(见 13.21 节和 15.13 节). 且任意特定样本值 $P = \overline{p}$ 是 p 的点估计值.

15.2　χ^2 统计量和 χ^2 分布

本章的主要目的是由包含在正态分布的单个随机样本中的信息得到总体方差 σ^2 的置信区间，为得到该参数的置信界，需要知道：(1) 样本方差 S^2 的点估计值 s^2，(2) 使 S^2 与 σ^2 有关的样本统计量，(3)该统计量的抽样分布.

就此目的常用的样本统计量是 χ^2 **统计量**：

$$X^2 = \frac{(n-1)S^2}{\sigma^2} \tag{15.1}$$

其中 X 是希腊字母 chi 的大写，记号 X^2 读作"卡方". 该统计量的特定值为

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \tag{15.2}$$

其中 χ 是 chi 的小写，χ^2 也读作"卡方".

χ^2 统计量是一随机变量，其抽样分布称为**卡方分布(或 χ^2 分布)**. 由抽样理论知：如果大小为 n 的所有可能的样本取自方差为 σ^2 的正态分布总体，且对每一样本计算随机变量 $X^2 = \frac{(n-1)S^2}{\sigma^2}$ 的特定值 $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$，则 χ^2 值将有一个称作卡方分布的连续概率分布(抽样分布).

与以前的连续概率分布(见 12.2，12.5，12.16 和 14.15 节)一样，χ^2 分布由一个特定的唯一的概率密度函数所定义. 略去推导，给出这一函数

$$f(\chi^2) = c(\chi^2)^{(\nu/2)-1}e^{-\chi^2/2} \tag{15.3}$$

其中 $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$

$e = 2.71828\cdots$(见上册习题 1.23)

$\nu = n-1$

c = 常数(为 ν 的函数)，使得卡方分布下总面积为 1.

这一函数具有与 t 分布一样的单参数 ν(符号可记作 d,df,d.f.,DF,r 或其它),也称为**自由度的个数**或简称**自由度**(见 14.15 和 14.17 节).与 t 分布一样,由于计算 χ^2 中有一个限制,即样本均值 $\bar{x}$ 的计算,故 $\nu = n-1$.如 14.17 节所指,由于概率分布是相互联系的,所以参数 ν 将在本书中经常出现.

与以前的连续概率分布一样,卡方分布不仅仅只是一个分布而是一个连续分布族,统称为卡方分布.对每一个整数值自由度($\nu = n-1$)都有唯一分布,在这无限多个分布中,图 15-1 给出了当 $\nu = 1,3,5,10$ 时的密度函数曲线.

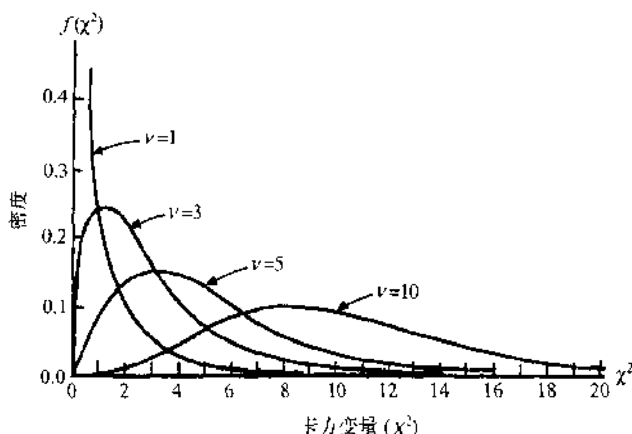


图 15-1

由图 15-1 可见,卡方分布与 t 分布和标准正态分布有很大的不同.后两种分布是关于零均值对称的,而卡方分布没有小于零的 χ^2 值,是不对称的(**正偏**).随着自由度在 $\nu = 2$ 以上增加,卡方分布变成单峰的,且越来越对称,但并不是关于零均值对称.事实上均值是自由度的个数,即 $\mu = \nu = n-1$.卡方分布的其它三个描述性参数是:方差 $\sigma^2 = 2\nu = 2(n-1)$;标准差 $\sigma = \sqrt{2\nu} = \sqrt{2(n-1)}$ 及(当 $\nu > 2$ 时)众数 $\nu - 2 = n - 3$.

15.3 χ^2 分布的临界值

在 12.10 节中,对连续正态分布随机变量 X 和标准正态变量 Z 给出了[(12.9)式]

$$P(X > x_\alpha) = P(Z > z_\alpha) = \alpha$$

而后,在 14.19 节中,对随机变量 T 给出了[(14.13)式]

$$P(T > t_{\alpha,\nu}) = \alpha$$

类似的,对于随机变量 X^2,现在我们给出

$$P(X^2 > \chi^2_{\alpha,\nu}) = \alpha \tag{15.4}$$

其中 $\chi^2_{\alpha,\nu}$ 记作右侧概率为 α,左侧概率为 $1-\alpha$ 的自由度为 ν 的 χ^2 分布临界值(或临界限).面积和临界值间的关系如图 15-2 所示.

标准正态分布和 t 分布是关于零均值对称的,所以在分布的正负尾部截断概率为 α 的临界值分别是:标准正态为 $-z_\alpha$ 和 z_α,t 分布为 $-t_{\alpha,\nu}$ 和 $t_{\alpha,\nu}$.相比之下,χ^2 分布并不关于均值对称,且当 $n>1$ 时均值不可能为零.因此,对于 χ^2 分布在负尾处左侧截断概率 α 的临界值为 $\chi^2_{1-\alpha,\nu}$,因此

$$P(X^2 < \chi^2_{1-\alpha,\nu}) = \alpha \tag{15.5}$$

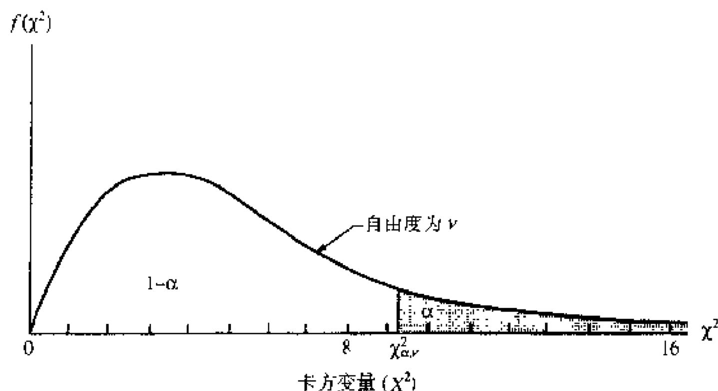


图 15-2

面积与临界值之间的关系如图 15-3 所示.

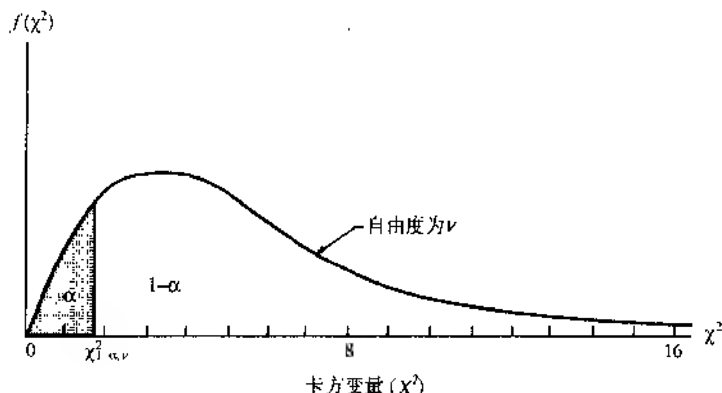


图 15-3

在 12.11 节中,对连续正态分布变量 X 和标准正态变量 Z 给出了[(12.18)式]

$$P(x_{1-\alpha/2} \leqslant X \leqslant x_{\alpha/2}) = P(-z_{\alpha/2} \leqslant Z \leqslant z_{\alpha/2}) = 1-\alpha$$

而后,在 14.19 中,对随机变量 T 给出了[(14.15)式]

$$P(-t_{\alpha/2,\nu} \leqslant T \leqslant t_{\alpha/2,\nu}) = 1-\alpha$$

现在,对于随机变量 X^2,我们给出

$$P(\chi^2_{1-\alpha/2,\nu} \leqslant X^2 \leqslant \chi^2_{\alpha/2,\nu}) = 1-\alpha \tag{15.6}$$

其中 $\chi^2_{\alpha/2,\nu}$ 表示自由度为 ν 的 χ^2 分布右尾截断 $\alpha/2$ 面积的临界值;$\chi^2_{1-\alpha/2,\nu}$ 表示同一 χ^2 分布左尾截断 $\alpha/2$ 面积的临界值.这些面积与临界值的关系如图 15-4 所示.

15.4　表 A.7:χ^2 分布的临界值

附录中表 A.7(χ^2 **分布的临界值**)给出了所选的 χ^2 分布($\nu=1,\cdots,30,40,50,60,70,80,90,100$)的临界值 $\chi^2_{\alpha,\nu}$.表中第一行代表不同的 χ^2 分布,左边一列标明了自由度 ν.每行中的值为给定的 χ^2 分布在其右侧截断特定面积 α 的临界值 $\chi^2_{\alpha,\nu}$,这个面积就是每列最上面的符号 χ^2_α 中的 α.表 15.1 为表 A.7 的一部分,给出了 $\nu=6,7,8,9,10$ 时的临界值.第一列给出右侧截断 $\alpha=0.995$ 的 $\chi^2_{0.995}$ 的值,最后一列给出右侧截断 $\alpha=0.005$ 的 $\chi^2_{0.005}$ 的值.

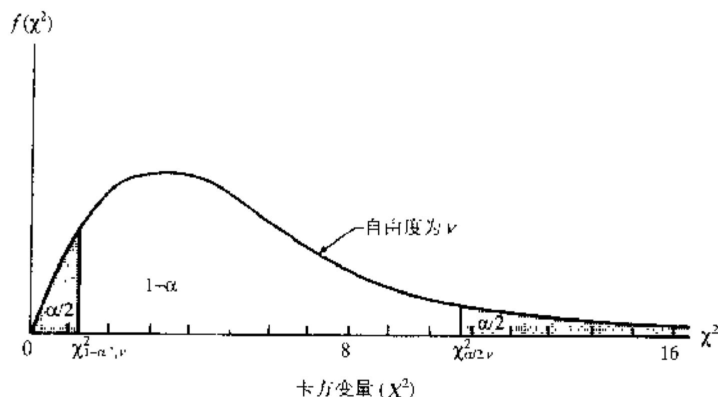


图 15-4

表 15.1*

ν	χ^2_α										
	$\chi^2_{0.995}$	$\chi^2_{0.990}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.900}$	$\chi^2_{0.500}$	$\chi^2_{0.100}$	$\chi^2_{0.050}$	$\chi^2_{0.025}$	$\chi^2_{0.010}$	$\chi^2_{0.005}$
6	0.676	0.872	1.24	1.64	2.20	5.35	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19

例 15.1　由表 15.1 求 $\chi^2_{\alpha,\nu}=\chi^2_{0.050,10}$ 和 $\chi^2_{1-\alpha,\nu}-\chi^2_{1-0.050,10}$ 的值.

解 ☞　为得到 $\chi^2_{0.050,10}$,我们用 $\chi^2_{0.050}$ 列和 $\nu=10$ 行,正确值为 18.31,在其右侧截断 $\alpha=0.05$ 的面积.因表中只给出 χ^2_α 的值,所以为得到 $\chi^2_{1-\alpha}$,首先应将它变换为等价的 χ^2_α,所以欲求 $\chi^2_{1-0.050,10}$,我们应先找到 $\nu=10$ 时的 $\chi^2_{0.950}$ 值.由 $\chi^2_{0.950}$ 列和 $\nu=10$ 行,得到 3.94,在其右侧截断面积 $\alpha=0.950$(等价于左侧的 0.050).

对一给定的 α 值,$\chi^2_{\alpha,\nu}$ 随 ν 增大而增大,由 A.7 中各列值可以看出临界值与自由度间的这种关系.这是因为随着 ν 的增加,分布的均值($\mu=\nu$)和标准差($\sigma=\sqrt{2\nu}$)都相应增加,且分布集中转换到右侧(见图 15-1).

对一给定的 ν 值,$\chi^2_{\alpha,\nu}$ 随 α 减小而增大,由对比表 A.7 中各行值可以看出临界值与 α 间的这种关系.这是因为:由于 α 定义为一个较小面积,因而在右侧截断 α 的临界值越来越向右.

15.5　正态分布总体方差 σ^2 的置信区间

如果所有大小为 n 的可能随机样本($X_1,X_2,\cdots,X_n$)来自一个无限大的正态分布总体,且对每一样本计算连续随机变量 S^2 的值 s^2,则有

$$P\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2,\nu}}\leqslant\sigma^2\leqslant\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,\nu}}\right]=1-\alpha \tag{15.7}$$

为证明上式的正确性,首先由(15.6)式

$$P(\chi^2_{1-\alpha/2,\nu}\leqslant X^2\leqslant\chi^2_{\alpha/2,\nu})=1-\alpha$$

将随机变量 $X^2=\dfrac{(n-1)S^2}{\sigma^2}$[(15.1)式]代入上式得:

* 表文所圈值有误,应为 3.94 和 18.31.——译者注

$$P\left[\chi^2_{1-\alpha/2,\nu} \leqslant \frac{(n-1)S^2}{\sigma^2} \leqslant \chi^2_{\alpha/2,\nu}\right] = 1-\alpha$$

不等式每一项分别除以$(n-1)S^2$,不改变不等号方向(见第一卷 1.23 节):

$$P\left[\frac{\chi^2_{1-\alpha/2,\nu}}{(n-1)S^2} \leqslant \frac{1}{\sigma^2} \leqslant \frac{\chi^2_{\alpha/2,\nu}}{(n-1)S^2}\right] = 1-\alpha$$

最后,各项取倒数并改变不等号方向得

$$P\left[\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,\nu}} \geqslant \sigma^2 \geqslant \frac{(n-1)S^2}{\chi^2_{\alpha/2,\nu}}\right] = 1-\alpha$$

这也可写成

$$P\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2,\nu}} \leqslant \sigma^2 \leqslant \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,\nu}}\right] = 1-\alpha$$

因而,如果对每一样本计算随机区间的特定值

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2,\nu}} \leqslant \sigma^2 \leqslant \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,\nu}}$$

则这些特定值的$(1-\alpha)100\%$

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}} \leqslant \sigma^2 \leqslant \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}} \tag{15.8}$$

将包含σ^2.因此称任意此种特定区间为σ^2的一个$(1-\alpha)100\%$**置信区间**.

这些求σ^2的$(1-\alpha)100\%$置信区间的方法只适用于正态分布总体,对偏离正态性是不稳健的.

15.6　置信限的表示

在 14.10 节中我们指出总体均值μ(σ已知,正态分布)的$(1-\alpha)100\%$置信区间可由下列形式之一表示:

(1) $\bar{x}-z_{\alpha/2}\sigma_{\bar{x}} \leqslant \mu \leqslant \bar{x}+z_{\alpha/2}\sigma_{\bar{x}}$,

(2) $L=\bar{x}-z_{\alpha/2}\sigma_{\bar{x}}$, $U=\bar{x}+z_{\alpha/2}\sigma_{\bar{x}}$,

(3) $(\bar{x}-z_{\alpha/2}\sigma_{\bar{x}}, \bar{x}+z_{\alpha/2}\sigma_{\bar{x}})$,

(4) $\bar{x}\pm z_{\alpha/2}\sigma_{\bar{x}}$.

而总体方差σ^2的$(1-\alpha)100\%$置信区间只能由下列三种形式表示:

(1) $\dfrac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}} \leqslant \sigma^2 \leqslant \dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}$

(2) $L=\dfrac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}}$, $U=\dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}$

(3) $\left[\dfrac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}}, \dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}\right]$

第四种形式之所以不能用是因为:

(1) Z统计量($Z=\dfrac{\bar{X}-\mu}{\sigma_{\bar{X}}}$)中点估计值与其参数是减法关系,而$\chi^2$统计量$\left[X^2=\dfrac{(n-1)S^2}{\sigma^2}\right]$则是比$\left(\dfrac{S^2}{\sigma^2}\right)$的关系;

(2) Z方法构造置信区间是由点估计值加减一个误差因子($\bar{x}\pm z_{\alpha/2}\sigma_{\bar{x}}$)得到的,而$\chi^2$方

法构造置信区间是由点估计值乘以一个误差因子$\left[\frac{(n-1)}{\chi^2}\times s^2\right]$得到的；

(3) 标准正态分布关于零均值对称，因而有对称形式($\bar{x}\pm z_{\alpha/2}\sigma_{\bar{x}}$)，而 χ^2 分布是不对称的.

例 15.2　在例 14.12 中，对于取自正态分布总体(μ,σ 未知)的随机样本($n=16,\bar{x}=27.9,s=3.23$)，我们得到总体均值 μ 的 95% 置信区间为 27.9 ± 1.72. 求总体方差 σ^2 的(a)95% 置信区间，(b)99% 置信区间？

解　(a)我们需要求

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}},\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}\right]$$

将下列特定值代入上式：

$$s^2=(3.23)^2=10.4329$$

$$\nu=n-1=16-1=15$$

$$\alpha/2=0.025$$

$$1-\alpha/2=0.975$$

由表 A.7 有

$$\chi^2_{\alpha/2,\nu}=\chi^2_{0.025,15}=27.49$$

$$\chi^2_{1-\alpha/2,\nu}=\chi^2_{0.975,15}=6.26$$

故 σ^2 的 95% 置信区间为

$$\left(\frac{15\times10.4329}{27.49},\frac{15\times10.4329}{6.26}\right)=(5.69,25.00)$$

(b) s^2 和 ν 的值同上，但 $\alpha/2=0.005$，$1-\alpha/2=0.995$，因此由表 A.7 有：

$$\chi^2_{\alpha/2,\nu}=\chi^2_{0.005,15}=32.80$$

$$\chi^2_{1-\alpha/2,\nu}=\chi^2_{0.995,15}=4.60$$

故 σ^2 的 99% 置信区间为

$$\left(\frac{15\times10.4329}{32.80},\frac{15\times10.4329}{4.60}\right)=(4.77,34.02)$$

15.7　方差置信区间的精度

由于总体均值 μ 的置信区间估计值可写成对称形式 $\bar{x}\pm$(误差边缘)，所以任一特定区间估计值的精度可由误差边缘[$E=z_{\alpha/2}\sigma_{\bar{x}}$(见 14.11 节)，$E=t_{\alpha/2,\nu}s_{\bar{x}}$(见 14.22 节)]来度量. 而对 σ^2 的置信区间估计值，却不能计算这样的 E 值(见 15.6 节)，然而我们可以回到原始的定义(见 14.11 节)，用特定区间估计的长度作为其估计值的精度的度量，即：

$$精度=(上界值)-(下界值)\tag{15.9}$$

例 15.3　就例 15.2 中的样本，分别给出：(a)95% 的置信区间，(b)99% 的置信区间的精度.

解　(a)对于 95% 的区间，精度 $=25.00-5.69=19.31$.

(b)对于 99% 的区间，精度 $=34.02-4.77=29.25$.

此例表明对给定样本的 σ^2 的区间估计的精度随着 $(1-\alpha)100\%$ 的增大而减小(区间长度相应增大).

例 15.4　如果将 σ^2 固定为常数 10.4329，那么随着样本量由 $n=16$ 增加至：(a) $n=31$，(b) $n=61$，例 15.2 中的 95% 置信区间的精度将如何变化？

解　(a) 当 $n=31$ 时，

$$s^2=10.4329$$

$$\nu=n-1=31-1=30$$

又由表 A.7 有

$$\chi^2_{\alpha/2,\nu} = \chi^2_{0.025,30} = 46.98$$

$$\chi^2_{1-\alpha/2,\nu} = \chi^2_{0.975,30} = 16.79$$

则 σ^2 的 95% 置信区间为：

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}\right] = \left(\frac{30 \times 10.4329}{46.98}, \frac{30 \times 10.4329}{16.79}\right) = (6.66, 18.64)$$

因此，对这个区间，

$$精度 = 18.64 - 6.66 = 11.98$$

(b) 当 $n = 61$ 时，

$$s^2 = 10.4329$$

$$\nu = n - 1 = 61 - 1 = 60$$

又由表 A.7 有：

$$\chi^2_{\alpha/2,\nu} = \chi^2_{0.025,60} = 83.30$$

$$\chi^2_{1-\alpha/2,\nu} = \chi^2_{0.975,60} = 40.48$$

则 σ^2 的 95% 置信区间为

$$\left(\frac{60 \times 10.4329}{83.30}, \frac{60 \times 10.4329}{40.48}\right) = (7.51, 15.46)$$

因此，对这个区间，

$$精度 = 15.46 - 7.51 = 7.95$$

此例表明：随着样本容量 n 的增加，σ^2 的置信区间的精度增加（区间长减小）.

15.8　确定为得到方差的所要求估计性质所必需的样本容量

我们在第十四章介绍了一些如何事先确定所需的样本容量，以得到 μ 的置信区间的所要求的估计值（见 14.12 和 14.22 节）. 这些方法给出了求解关于 n 的方程[E =（误差边缘）]，遗憾的是，对 σ^2 的置信区间有许多预先确定样本容量的方法，但并非显而易见.

因为精确的 χ^2 方法不能产生形如 $s^2 \pm$（误差边缘）的对称置信区间，所以不可能得到一个关于 n 精确的方程. 然而如果样本容量（从前的及未来的）都很大（$n \geq 100$），则用 S^2 的抽样分布的正态近似（见 15.10 节）可得到一个预先确定 n 的近似方法. 除此以外若给定一个 s^2 的从前点估计值，并已知所要求的估计性质（置信水平和大致的所要求区间长度），则对不同样本容量计算一系列置信区间（见例 15.4）直到达到所要求的性质. 并且还可以同时估计 μ 和 σ^2，来确定估计 μ 所需的样本容量.

15.9　用近似正态方法确定方差的置信区间

由图 15-1 可知，随着自由度 ν 的增加，χ^2 分布越来越对称，而且在数学上可以证明，当 ν 足够大时，χ^2 分布接近正态分布. 可惜的是，实际应用中没有那么大的样本允许 χ^2 分布近似正态.

但对于一个比较合理的样本容量[$(\nu = n - 1) > 30$]，可以证明随机变量 $\sqrt{2X^2} - \sqrt{2\nu - 1}$ 的抽样分布基本上是标准正态分布（$\mu = 0, \sigma^2 = \sigma = 1$）. 于是，当 $\nu > 30$ 时有

$$Z \approx \sqrt{2X^2} - \sqrt{2\nu - 1}$$

移项后得

$$X^2 \approx \frac{1}{2}(Z + \sqrt{2\nu - 1})^2 \tag{15.10}$$

当 $\nu>30$ 时，为确定 σ^2 的置信区间，可用此等式计算近似的 χ^2 值.

例 15.5 在例 15.4 中，我们知道当 $\alpha=0.05, n=61$ 时有：

$$\chi^2_{\alpha/2,\nu}=\chi^2_{0.025,60}=83.30$$

$$\chi^2_{1-\alpha/2,\nu}=\chi^2_{0.975,60}=40.48$$

试用上述近似方法得到这些值.

解 ☞ 在题设条件下，可以用(15.10)式，得到

$$\chi^2_{\alpha/2,\nu}\approx\frac{1}{2}(z_{\alpha/2}+\sqrt{2\nu-1})^2$$

和

$$\chi^2_{1-\alpha/2,\nu}\approx\frac{1}{2}(-z_{\alpha/2}+\sqrt{2\nu-1})^2$$

当 $\alpha=0.05$ 时，由表 A.5 知 $z_{\alpha/2}=z_{0.025}=1.96$，于是 $-z_{\alpha/2}=-z_{0.025}=-1.96$. 将这些值及 $\nu=n-1=61-1=60$ 代入上式，则有

$$\chi^2_{0.025,60}\approx\frac{1}{2}[z_{0.025}+\sqrt{(2\times60)-1}]^2=\frac{1}{2}(1.96+\sqrt{119})^2=82.80$$

$$\chi^2_{0.975,60}\approx\frac{1}{2}[-z_{0.025}+\sqrt{(2\times60)-1}]^2=\frac{1}{2}(-1.96+\sqrt{119})^2=40.04$$

将现结果与例 15.4 中的结果(83.30,40.48)比较可知，当 $\nu>30$ 时，这一近似正态给出了 χ^2 值的合理近似值.

15.10 用样本方差的抽样分布来近似总体方差的置信区间

如果来自正态分布总体的样本很大($n\geqslant100$)，则可用随机变量 S^2(样本方差)的抽样分布来近似得到 σ^2 的$(1-\alpha)100\%$置信区间. 不考虑数学推导，我们简单叙述如下：

如果样本很大($n\geqslant100$)且来自正态分布总体，则 S^2 的抽样分布基本是正态的，均值 $\mu_{s^2}=\sigma^2$、标准差(标准误差) $\sigma_{s^2}=\sigma^2\sqrt{\frac{2}{n}}$.

由上可知，在上述条件下，σ^2 的$(1-\alpha)100\%$的置信区间可由下式近似：

$$s^2\pm z_{\alpha/2}\sigma^2\sqrt{\frac{2}{n}} \tag{15.11}$$

用点估计值 s^2 来估计 σ^2，上式变为：

$$s^2\pm z_{\alpha/2}s^2\sqrt{\frac{2}{n}} \tag{15.12}$$

例 15.6 若例 15.4 中的样本，s^2 仍为 10.4329，但 n 增加为 350. 试用(15.12)式给出 σ^2 的近似 95%置信区间.

解 ☞ 由表 A.5 得 $z_{0.05/2}=1.96$，于是 σ^2 的近似 95%置信区间为

$$10.4329\pm\left(1.96\times10.4329\times\sqrt{\frac{2}{350}}\right)$$

$$10.43\pm1.55$$

或

$$(8.88, 11.98)$$

用这一近似置信区间的优点在于，它提供了一个近似误差边缘方程，因而对于(15.11)式有

$$E\approx z_{\alpha/2}\sigma^2\sqrt{\frac{2}{n}} \tag{15.13}$$

解出 n 得

$$n \approx \left(\frac{\sqrt{2}\, z_{\alpha/2}\sigma^2}{E}\right)^2$$

用 s^2 代换 σ^2,我们可由

$$n \approx \left(\frac{\sqrt{2}\, z_{\alpha/2}s^2}{E}\right)^2 \tag{15.14}$$

来确定一个近似的样本容量,以得到所要求的估计性质.

例 15.7　若想重复例 15.6 的方法得到 $E=1.00$,而不是 $E=1.55$ 时的近似 95% 置信区间,当 $s^2=10.4329$ 时,用(15.14)式给出所需样本容量.

解　将 $E=1.00$,　$z_{0.05/2}=1.96$,　$s^2=10.4329$ 代入方程得

$$n \approx \left(\frac{\sqrt{2}\times 1.96\times 10.4329}{1.00}\right)^2 = 836.28 \text{ 或 } 837$$

因此,若使误差边缘 $E=1.55$ 减小到 $E=1.00$,需将样本容量由 350 至少增加至 837.

15.11　正态分布总体标准差 σ 的置信区间

通过对方差 σ^2 的置信区间(15.6)式上下限取算数平方根,可算出总体标准差 σ 的精确的 $(1-\alpha)100\%$ 置信区间:

$$\left[\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}}\right] \tag{15.15}$$

例 15.8　在去医院进行身体检查前,你测量了 21 次自己的血压.这 21 个舒张压(mmHg)为 77,85,84,81,82,78,81,79,84,83,80,80,83,81,79,81,80,79,82,78,82.首先,用上册习题 7.19 中的频数分布方法来得到样本标准差 s,然后在总体假设为正态分布条件下给出舒张压测量值总体的 σ 的精确的 95% 置信区间.

解　表 15.2 的下面给出了样本标准差:$s=2.16576$mmHg,于是 $s^2=4.69052$mmHg^2.又知 $n-1=21-1=20$,由附录表 A.7 可得

$$\chi^2_{\alpha/2,\nu} = \chi^2_{0.05/2,20} = 34.17$$

$$\chi^2_{1-\alpha/2,\nu} = \chi^2_{1-0.05/2,20} = 9.59$$

表 15.2

舒张压 (mmHg) x_i	频数 f_i	$f_i x_i$ (mmHg)	$(x_i-\bar{x})$ (mmHg)	$(x_i-\bar{x})^2$ (mmHg^2)	$f_i(x_i-\bar{x})^2$ (mmHg^2)
77	1	77	−3.9	15.21	15.21
78	2	156	−2.9	8.41	16.82
79	3	237	−1.9	3.61	10.83
80	3	240	−0.9	0.81	2.43
81	4	324	0.1	0.01	0.04
82	3	246	1.1	1.21	3.63
83	2	166	2.1	4.41	8.82
84	2	168	3.1	9.61	19.22
85	1	85	4.1	16.81	16.81
$\sum$	21	1,699mmHg			93.81mmHg^2

$$\bar{x} = \frac{\sum f_i x_i}{n} = \frac{1{,}699\text{mmHg}}{21} = 80.9\text{mmHg}$$

$$s = \sqrt{\frac{\sum f_i(x_i-\bar{x})^2}{n-1}} = \sqrt{\frac{93.81\text{mmHg}^2}{20}} = 2.16576\text{mmHg}$$

将这些值代入(15.15)式得到 σ 的精确的95%的置信区间为

$$\left(\sqrt{\frac{20\times 4.69052\text{mmHg}^2}{34.17}},\sqrt{\frac{20\times 4.69052\text{mmHg}^2}{9.59}}\right)=(1.66\text{mmHg},3.13\text{mmHg})$$

15.12 用样本标准差的抽样分布来近似总体标准差的置信区间

总体标准差 σ 的 $(1-\alpha)100\%$ 置信区间可由随机变量 S(样本标准差)的抽样分布近似得到,类似于15.10节中的方法,即用样本方差 S^2 的抽样分布来近似总体方差 σ^2 的置信区间[(15.12)式].仍不考虑数学推导,叙述如下:

如果样本很大($n\geqslant 100$)且来自正态分布总体,则 S 的抽样分布基本是正态的,均值 $\mu_s=\sigma$,标准差(标准误差)$\sigma_s=\dfrac{\sigma}{\sqrt{2n}}$.

由以上条件知 σ 的 $(1-\alpha)100\%$ 置信区间近似为

$$s\pm z_{\alpha/2}\frac{\sigma}{\sqrt{2n}}$$

用点估计值 s 来估计 σ,则任意特定近似置信区间为

$$s\pm z_{\alpha/2}\frac{s}{\sqrt{2n}} \tag{15.16}$$

例 15.9 在例15.6中对于来自正态分布总体的样本($n=350,s^2=10.4329$),我们用(15.12)式得到方差的近似95%置信区间为(8.88,11.98).用下面两种方法得到该总体 σ 的近似95%置信区间:(a)对近似边界值取算术平方根,(b)用(15.16)式求解.

解 ☞ (a) $(\sqrt{8.88},\sqrt{11.98})=(2.98,3.46)$;

(b) 将 $s=\sqrt{10.4329}=3.23$,$z_{\alpha/2}=z_{0.05/2}=1.96$ 及 $2n=2(350)=700$ 代入(15.16)式,得

$$3.23\pm\left(1.96\times\frac{3.23}{\sqrt{700}}\right)$$

$$3.23\pm 0.24$$

或(2.99,3.47).

15.13 二项总体比率 p 的最优估计

我们在13.12节指出,二项总体中成功比率 p 可由随机变量 $\bar{P}$ 估计,其中 $\bar{P}$ 是来自总体的随机样本的成功比率,该变量由公式[(13.22)式]定义为

$$\bar{P}=\frac{Y}{n}$$

其中 Y 是样本中成功的二项随机变量个数,n 为样本容量.因为 $\bar{P}$ 抽样分布的均值为 p [(13.27)式]:$E(\bar{P})=\mu_{\bar{P}}=p$,所以 $\bar{P}$ 是 p 的无偏估计.鉴于这一原因及其它因素认为,$\bar{P}$ 是 p 的最优估计量.因此对一给定样本,由样本计算出特定的 $\bar{P}=\bar{p}$ 值是 p 的最优点估计值.

例 15.10 假设一个无限大的二项总体具有未知的成功元素比率 p,在Bernoulli试验条件下,从总体选取 $n=150$ 个元素的随机样本(见11.2节),知样本中有 $y=82$ 个成功元素,求参数 p 的最优点估计值 $\hat{\theta}^*$.

解 ☞

$$\bar{p}=\frac{y}{n}=\frac{82}{150}=0.546667\text{ 或 }0.55$$

15.14 二项总体比率 p 的近似置信区间的导出

我们已经在13.23节知道,如果在Bernoulli试验条件下,大小为 n 的所有可能随机样本

取自一无限大二项总体,其中二项总体的成功元素比率为 p,且对每一样本计算样本比率 $\bar{P} = \bar{p}$. 则比率的理论抽样分布是均值[(13.27)式]$\mu_{\bar{p}} = p$,标准差(比率的标准误差)[(13.29)式]

$$\sigma_{\bar{p}} = \sqrt{\frac{pq}{n}}$$

的二项分布,其中 $q = 1 - p$. 由 13.25 节知,如果 n"足够大"($np \geqslant 5$ 且 $nq \geqslant 5$),则比率的抽样分布是 $\mu_{\bar{p}} = p$, $\sigma_{\bar{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}$ 的近似正态分布. 因此,在此条件下随机变量[(13.30)式]

$$Z = \frac{\bar{P} - \mu_{\bar{p}}}{\sigma_{\bar{p}}} = \frac{\bar{P} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

将有近似标准正态分布. 由此可得到**二项比率 p 的近似$(1-\alpha)100\%$置信区间**的表达式.

由(12.18)式有

$$P(-z_{\alpha/2} \leqslant Z \leqslant z_{\alpha/2}) = 1 - \alpha$$

用(13.30)式代换 Z 有

$$P\left[-z_{\alpha/2} \leqslant \frac{\bar{P} - p}{\sqrt{\frac{p(1-p)}{n}}} \leqslant z_{\alpha/2}\right] \approx 1 - \alpha$$

不等式各项分别乘以$\sqrt{\frac{p(1-p)}{n}}$并减去$\bar{P}$得

$$P\left[-\bar{P} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leqslant -p \leqslant -\bar{P} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right] \approx 1 - \alpha$$

最后,不等式各项乘以 -1 得

$$P\left[\bar{P} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \geqslant p \geqslant \bar{P} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right] \approx 1 - \alpha$$

上式也可写作

$$P\left[\bar{P} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leqslant p \leqslant \bar{P} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right] \approx 1 - \alpha \tag{15.17}$$

对于 Bernoulli 试验条件下取自一无限大二项总体的"足够大"样本,如果对所有大小为 n 的样本计算随机区间的特定值

$$\bar{P} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leqslant p \leqslant \bar{P} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

则这些样本区间

$$\bar{p} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leqslant p \leqslant \bar{p} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \tag{15.18}$$

中约有$(1-\alpha)100\%$包含 p. 因此,称这种特定区间为二项比率 p 的近似$(1-\alpha)100\%$置信区间.

这一近似置信区间可表示为下列形式之一(见 14.10 节):

(1) $\bar{p} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leqslant p \leqslant \bar{p} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$,

(2) $L=p-z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$, $U=\bar{p}+z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$,

(3) $\left[\bar{p}-z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}, p+z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right]$,

(4) $\bar{p}\pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$.

15.15　参数 p 的估计

在 15.14 节中,我们用近似正态的方法得到了二项比率 p 的近似$(1-\alpha)100\%$置信区间,即 $\bar{p}\pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$.但在该置信区间上有一个明显的问题,即对一给定样本计算置信限必须首先知道待估的总体参数 p;而解决 p 未知的问题有两种方法:**点估计方法**和**保守解法**.

在点估计方法中,如果 n"足够大",则样本点估计值 $\bar{p}$ 可以代替置信限中的参数 p.于是置信限可以写作

$$\bar{p}\pm z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{15.19}$$

保守解法基于下列事实:不顾 p 的取值,在数学上可以证明 $p(1-p)$ 的最大可能值为 $\frac{1}{4}$(即当 $p=0.5$ 且 $1-p=0.5$ 时).因此如果 n 不是"足够大",或者取更多的保守解时,可用$\frac{1}{4}$来替换 $p(1-p)$,于是置信区间可写作

$$\bar{p}\pm z_{\alpha/2}\sqrt{\frac{1/4}{n}}$$

或

$$p\pm z_{\alpha/2}\sqrt{\frac{1}{4n}} \tag{15.20}$$

保守解给出一个最大可能长度的置信区间,因而可以说此区间的置信水平至少为$(1-\alpha)100\%$.

例 15.11　就例 15.10 中的样本,用近似正态方法分别由(a)点估计值和(b)保守解计算未知的总体比率 p 的近似 95%置信区间.

解 ☞　(a) 将 $n=150$, $p=0.546667$, $z_{\alpha/2}=z_{0.05/2}=1.96$ 代入(15.19)式,得

$$0.546667\pm\left[1.96\times\sqrt{\frac{0.546667(1-0.546667)}{150}}\right]$$

$$0.55\pm0.0797$$

(b) 将 $n=150$, $\bar{p}=0.546667$ 代入(15.20)式,得

$$0.546667\pm\left(1.96\times\sqrt{\frac{1}{4\times150}}\right)$$

$$0.55\pm0.0800$$

正如我们所期望的,保守解的误差边缘(0.0800)比点估计方法的(0.0797)要稍微大一些.

15.16　当 p 未知时,确定何时 n 为"足够大"

当用正态近似比率的抽样分布时,我们说 $np\geqslant5$, $nq\geqslant5$ 同时成立时 n 是"足够大"的(见

15.14 节). 但这是在假定总体参数 p 已知,且是用一个分布近似另一个分布的. 当 p 未知时,为计算 p 的置信区间,不但需要用正态分布来近似二项分布,而且还要用样本估计值 $\bar{p}$ 或 $p(1-p)$ 的最大可能值来近似 p(见 15.15 节).

对于未知 p,统计著作中对 $np \geqslant 5$ 且 $nq \geqslant 5$ 时 n 是否为"足够大"持有不同意见. 一些著作中认为足够大,而大部分著作中建议应谨慎对待,给出的更保守的"足够大"n 值条件有:$np \geqslant 10$ 且 $nq \geqslant 10$, $np \geqslant 15$ 且 $nq \geqslant 15$ 及 $n \geqslant 30$[该条件同于应用中心极限定理的限制(见 13.18 节)],且更为保守的是仅当 $n \geqslant 100$ 时为足够大.

15.17　有限总体无放回抽样对二项参数 p 的近似置信区间

15.14 和 15.15 节中近似置信区间公式是在假定抽样的二项总体无限大时得到的,这些公式对有限二项总体的有放回抽样也成立,但对有限二项总体的无放回抽样却不成立. 如果 $n > 0.05N$,则点估计法和保守解(见 15.15 节)应含有有限总体修正因子(见 13.11 节),即成为

$$\bar{p} \pm z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}\sqrt{\frac{N-n}{N-1}} \tag{15.21}$$

和

$$\bar{p} \pm z_{\alpha/2}\sqrt{\frac{1}{4n}}\sqrt{\frac{N-n}{N-1}} \tag{15.22}$$

在这些条件下,当 $n \leqslant 0.05N$ 时,不需要有限总体修正因子.

15.18　二项参数 p 的精确的置信区间

所有用近似正态方法(见 15.14,15.15 和 15.17 节)得到的二项参数 p 的 $(1-\alpha)100\%$ 置信区间都是近似的表达式. 也存在精确的方法,其中包括了二项分布(见 11.5 节)和超几何分布(见 11.21 节)方法,但它们都很复杂而很少在统计著作中予以介绍.

15.19　二项参数 p 的近似置信区间估计的精度

在确定 σ 已知时总体均值的置信区间的精度时(见 14.11 节),我们指出对于任何特定置信区间估计 $\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$,估计的精度是由误差边缘 $E = z_{\alpha/2}\sigma_{\bar{x}}$ 度量的(误差边缘越小,估计的精度越高). 类似地,二项参数 p 的**近似置信区间估计**(见 15.14 节)

$$\bar{p} \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

的精度也是用近似误差边缘

$$E \approx z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \tag{15.23}$$

来度量的. 由于 p 未知,15.15 节知给出了 p 的近似 $(1-\alpha)100\%$ 置信区间,若用点估计方法[(15.19)式],其精度为

$$E \approx z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}} \tag{15.24}$$

若用保守解[(15.20)式],其精度为

$$E \approx z_{\alpha/2}\sqrt{\frac{1}{4n}} \tag{15.25}$$

15.20　确定二项参数 p 近似置信区间的样本容量

在总体均值置信区间的研究中，我们发现所需样本容量依赖于所要求的估计性质：精度和置信水平(见 14.12 节)．如果 σ 已知，在均值的误差边缘表达式：$E=z_{\alpha/2}\sigma_{\bar{x}}=z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ 中可找到置信区间的两个性质．解方程求 n，即是在特定置信水平$(1-\alpha)$及精度(E)下所需的样本容量[(14.8)式]．类似地，对于二项参数 p 的近似$(1-\alpha)100\%$置信区间估计，

$$E\approx z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

两个性质—置信水平和精度—都出现在误差边缘中．解方程求 n，得到所需样本容量：

$$n\approx\left[\frac{z_{\alpha/2}\sqrt{p(1-p)}}{E}\right]^2 \tag{15.26}$$

实际上，由于 p 未知，我们可用 15.15 节中的一个解来确定样本容量 n．对于点估计法[(15.19)式]，误差边缘为

$$E\approx z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

解方程求 n，所需样本容量为

$$n\approx\left[\frac{z_{\alpha/2}\sqrt{\bar{p}(1-\bar{p})}}{E}\right]^2 \tag{15.27}$$

对于保守解[(15.20)式]，误差边缘为

$$E\approx z_{\alpha/2}\sqrt{\frac{1}{4n}}$$

解方程求 n，所需样本容量为

$$n\approx\left(\frac{z_{\alpha/2}}{2E}\right)^2 \tag{15.28}$$

例 15.12　在例 15.11 中，我们对例 15.10 中的样本得到了 p 的近似 95%置信区间估计．并得到点估计法解 0.55 ± 0.0797 以及保守法解 0.55 ± 0.080．如果我们仍研究这个问题，想得到误差边缘为 $E=0.05$ 的近似 95%置信区间，那么对以上两个解，求所需样本容量．

解　将 $z_{\alpha/2}=z_{0.05/2}=1.96$，$E=0.05$ 及 $\bar{p}=0.546667$(来自例 15.10)代入(15.27)式，所需样本量为

$$n\approx\left[\frac{1.96\sqrt{0.546667(1-0.546667)}}{0.05}\right]^2=380.813 \text{ 或 } 381$$

同样将 $z_{0.05/2}$ 和 E 值代入(15.28)式，所需样本量为

$$n\approx\left(\frac{1.96}{2\times0.05}\right)^2=384.160 \text{ 或 } 385$$

如果恰好存在一个 $\bar{p}$ 的先验点估计值 $\bar{p}$，则可将该值用于点估计法[(15.27)式]来确定重复研究所需样本量．如果存在多个的先验 $\bar{p}$ 值，则给出如下说明：

(1) 如果先验 $\bar{p}$ 值的范围全小于 0.5 或全大于 0.5，则选择接近于 0.5 的值，再由(15.27)式计算 n 值．

(2) 如果先验 $\bar{p}$ 值的范围含有 0.5，则直接用(15.28)式计算 n 值．

例 15.13　习题 14.5 为得到所要求平均生长时间精度水平，给出了快速生长的洋葱所需样本容量．现在研究发芽种子的比率 p．从以前的研究得到三个点估计值 $\bar{p}$：0.75，0.78，0.85．如今想重复此研究得到 p 的近似 95%置信区间，要求误差边缘 $E=0.02$ 或更小．问试验中需要种子的样本容量 n 为多少？

解 ☞ 由说明(1),将 $p=0.75$, $z_{0.05/2}=1.96$ 及 $E=0.02$ 代入(15.27)式得

$$n \approx \left[\frac{1.96\sqrt{0.75(1-0.75)}}{0.02}\right]^2 = 1,800.75 \text{ 或 } 1,801$$

计算样本容量 n 的两个公式[(15.27)和(15.28)式]是基于对一个无限大总体或至少是很大的总体(如 $N>25000$)抽样的假定. 如果总体没有那么大,则两个公式都给出了过大的 n 值. 为解决这一问题,推荐一种两阶段方法:

(1) 首先用(15.27)或(15.28)式来计算 n 值;

(2) 然后对小总体用下式修正 n:

$$\text{修正样本量} \approx \frac{n}{1+\frac{n}{N}} \tag{15.29}$$

其中 n 是第一步中确定的样本容量, N 是抽样总体的容量.

15.21 二项总体百分比的近似置信区间

得到二项总体比率 p 的 $(1-\alpha)100\%$ 近似置信区间后,在区间端点上乘以 100% 则可算出百分比 $p\times100\%$ 的近似置信区间. 通常将百分比置信区间用于投票选举结果中(见习题15.18至 15.28).

例 15.14 在快速生长的洋葱研究中(见例 15.13),在所取的 1000 颗种子的样本中发现有 760 粒发芽($\bar{p}=0.76$). 假定样本足够大,且在 Bernoulli 试验条件下,给出发芽的总体百分比 $p\times100\%$ 的近似 95% 置信区间.

解 ☞ 首先将 $\bar{p}=0.760$, $z_{0.05/2}=1.96$ 及 $n=1000$ 代入(15.19)式,得到总体比率的近似 95% 置信区间为

$$0.76 \pm 1.96\sqrt{\frac{0.76(1-0.76)}{1,000}}$$

$$0.76 \pm 0.026$$

而后将该区间端点同时乘以 100%,得到总体百分比的近似 95%置信区间为

$$(0.76\times100\%) \pm (0.026\times100\%)$$

$$76\% \pm 2.6\%$$

15.22 二项总体总的成功次数的近似置信区间

对区间端点同乘以总体大小 N,即可由二项总体比率 p 的近似置信区间变换为**总的"成功"总数**的置信区间.

例 15.15 例 15.14 给出了洋葱种子总体发芽比率 p 的近似 95% 置信区间,并就 1,000 颗种子的样本,得到了 p 的近似 95% 置信区间 0.76 ± 0.026. 若种子公司有 100,000 颗这种种子,该总体发芽种子总数的近似 95% 置信区间是什么?

解 ☞ 若要将比率 p 的置信区间变换为种子发芽总数的置信区间,我们将区间端点同时乘以总体容量(100,000)得

$$(0.76\times100,000) \pm (0.026\times100,000)$$
$$76,000 \pm 2,600$$

15.23 估计总体容量 N 的捕获-再捕获方法

由于总体太大,流动性太强,太隐蔽,太分散或其它原因而无法对总体容量计数时,则可用**捕获-再捕获方法**来估计总体大小. 该方法常常用于估计野生动物总体,但也可用于城市中吸毒上瘾的总体,捕获-再捕获方法分为两个步骤:

(1) 从总体中取一个样本,对每个元素做上标记,然后放回总体中,我们将该样本容量记

作 n_1；

(2) 经过"足够长的时间"后样本中的 n_1 个元素已分散在总体中，取容量为 n_2 的第二个样本，此样本中有 r 个来自第一样本在捕获(即有标记的)的元素.

我们假定在第二个样本中再捕获的元素比率：

$$\bar{p}=\frac{r}{n_2}$$

是总体中标记元素比率

$$p=\frac{n_1}{N}$$

的一个点估计值，则可以说

$$\frac{r}{n_2}\approx\frac{n_1}{N}$$

由此解出 N，则得到总体大小的近似估计值：

$$N\approx\frac{n_1 n_2}{r} \tag{15.30}$$

例 15.16 Minnesota 州想用捕获-再捕获方法来估计该地区狼群大小.首先在随机选择的捕猎处捉到 65 只狼，在其耳尖做了标记后放回；一个月后，再随机选择捕猎处，捉住 85 只狼，其中 3 只是有标记的(来自第一样本耳尖有标记的个体).问 Minnesota 州狼群个数 N 的近似值为多少?

解 ☞ 由(15.30)式，用捕获-再捕获方法估计 N，其中 $n_1=65, n_2=85, r=3$，则有

$$N\approx\frac{65\times 85}{3}=1{,}841.67 \text{ 或 } 1{,}842$$

习题解答

表 A.7：χ^2 分布的临界值

15.1 利用表 A.7 求 $\chi^2_{\alpha/2,\nu}=\chi^2_{0.100/2,10}$ 和 $\chi^2_{1-\alpha/2,\nu}=\chi^2_{1-0.100/2,10}$ 的值.

解 ☞ 与表 A.6 中 t 分布不同，表 A.7 没有直接给出 $\alpha/2$ 的临界值，首先需将 $\chi^2_{\alpha/2,\nu}$ 和 $\chi^2_{1-\alpha/2,\nu}$ 值转换成等价的 $\chi^2_{\alpha,\nu}$ 值，于是有

$$\chi^2_{\alpha/2,\nu}=\chi^2_{0.100/2,10}=\chi^2_{0.050,10}=18.31$$

和

$$\chi^2_{1-\alpha/2,\nu}=\chi^2_{1-0.100/2,10}=\chi^2_{1-0.050,10}=\chi^2_{0.950,10}=3.94$$

15.2 由表 A.7 给出下式各值：

(a) $\chi^2_{\alpha,\nu}=\chi^2_{0.005,27}$，(b) $\chi^2_{1-\alpha,\nu}=\chi^2_{1-0.025,17}$，(c) $\chi^2_{\alpha/2,\nu}=\chi^2_{0.010/2,100}$.

解 ☞ (a) $\chi^2_{\alpha,\nu}=\chi^2_{0.005,27}=49.64$

(b) $\chi^2_{1-\alpha,\nu}=\chi^2_{1-0.025,17}=\chi^2_{0.975,17}=7.56$

(c) $\chi^2_{\alpha/2,\nu}=\chi^2_{0.010/2,100}=\chi^2_{0.005,100}=140.2$

15.3 利用表 A.7 指出随着自由度的增加：5,10,30,50,100，$\chi^2_{\alpha,\nu}=\chi^2_{0.025,\nu}$ 是如何变化的?

解 ☞ $\chi^2_{0.025,5}=12.83$

$\chi^2_{0.025,10}=20.48$

$\chi^2_{0.025,30}=49.98$

$\chi^2_{0.025,50}=71.42$

$\chi^2_{0.025,100}=129.6$

15.4 由表 A.7 指出随着 α 的增加：0.100，0.050，0.025，0.010，0.005，$\chi^2_{\alpha,20}$ 是如何变化的?

解 ☞　$\chi^2_{0.100,20}=28.41$

$\chi^2_{0.050,20}=31.41$

$\chi^2_{0.025,20}=34.17$

$\chi^2_{0.010,20}=35.57$

$\chi^2_{0.005,20}=40.00$

方差的置信区间

15.5　若随机样本($n=16,\bar{x}=27.9$ 及 $s=3.23$)取自正态分布总体,求总体方差 σ^2 的 90%置信区间.

解 ☞　σ^2 的置信限为[15.6 节的第(3)种形式]:

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}},\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}\right]$$

其中

$$s^2=(3.23)^2=10.4329$$

$$\nu=n-1=16-1=15$$

于是当 $1-\alpha=0.90,\alpha=0.10,\alpha/2=0.05$ 及 $1-\alpha/2=0.95$ 时,由表 A.7 有

$$\chi^2_{\alpha/2,\nu}=\chi^2_{0.05,15}=25.00$$

$$\chi^2_{1-\alpha/2,\nu}=\chi^2_{0.95,15}=7.26$$

因此,σ^2 的 90%置信区间为

$$\left(\frac{15\times 10.4329}{25.00},\frac{15\times 10.4329}{7.26}\right)=(6.26,21.56)$$

15.6　由例 15.4 知当 $s^2=10.4329,n=61$ 时,σ^2 的 95%置信区间为(7.51,15.46).而在例 15.6 中,s^2 保持 10.4329 不变,n 增加至 350,由 S^2 的抽样分布知,σ^2 的近似 95%置信区间为(8.88,11.98).如今 $s^2=10.4329,n=350$ 不变,试用 15.9 节中正态近似方法,给出 σ^2 的近似 95%置信区间.

解 ☞　首先将 $z_{0.025}=1.96,-z_{0.025}=-1.96,s^2=10.4329,(\nu=n-1=350-1=349)$代入(15.10)式,得

$$\chi^2_{0.025,349}\approx\frac{1}{2}\left[z_{0.025}+\sqrt{(2\times 349)-1}\right]^2=\frac{1}{2}(1.96+\sqrt{697})^2=402.17$$

$$\chi^2_{0.975,349}\approx\frac{1}{2}\left[-z_{0.025}+\sqrt{(2\times 349)-1}\right]^2=\frac{1}{2}(-1.96+\sqrt{697})^2=298.68$$

然后将这些近似 χ^2 值代入置信限,得 σ^2 的近似 95%置信区间为:

$$\left(\frac{349\times 10.4329}{402.17},\frac{349\times 10.4329}{298.68}\right)=(9.05,12.19)$$

注意,此近似置信区间相当接近于例 15.6 中的近似区间.还注意到,随着 n 由 61 增至 350,精度[区间长度(见 15.7 节)]由(15.46−7.51=7.95)增至例 15.6 中的(11.98−8.88=3.10),而在此题中增至(12.19−9.05=3.14).

15.7　为研究学习过程中的记忆力,一位生理学家计划在迷宫试验之前或之后将激素注入老鼠体内.他首先研究 41 只未被注射的老鼠并且(见上册,3.10 节)记录每只老鼠能达到正确的跑出迷宫的标准之前,总的迷宫错误(反向)次数:$\bar{x}=13.2,s^2=5.19$.如果假定错误次数总体是正态分布的,确定总体的 μ 和 σ^2 的精确的 95%置信区间.

解 ☞　对于 μ,由 14.21 和 14.23 节知此条件下的置信区间为 $\bar{x}\pm t_{\alpha/2,\nu}s_{\bar{x}}$.对本题有

$$s_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{\sqrt{5.19}}{\sqrt{41}}=0.355788$$

并由表 A.6 知

$$t_{\alpha/2,\nu} = t_{0.05/2,40} = 2.021$$

因此,μ 的精确的95%置信区间为

$$13.2 \pm (2.021 \times 0.355788)$$

$$13.2 \pm 0.72$$

或

$$(12.48, 13.92)$$

为求得 σ^2 的精确的95%置信区间,将 $n-1=40$, $s^2=5.19$ 及由表 A.7 得到的

$$\chi^2_{\alpha/2,\nu} = \chi^2_{0.05/2,40} = 59.34$$

$$\chi^2_{1-\alpha/2,\nu} = \chi^2_{1-0.05/2,40} = 24.43$$

代入置信限中(见15.6节)得

$$\left(\frac{40 \times 5.19}{59.34}, \frac{40 \times 5.19}{24.43}\right) = (3.50, 8.50)$$

15.8 如果你在食品公司就职,要求确定一标准袋薯片的平均"总脂肪"量[单位:克].现分析了101袋,并得到下列结果:$\bar{x}=18.2g$, $s^2=0.56g$.如果假定总脂肪量总体是正态分布的,给出总体 μ 和 σ^2 的90%置信区间.

解 ☞　对于 μ,由于表 A.6 未给出 $\nu=n-1=100$ 时的 t 值,所以我们可以用近似置信区间(见14.23节)$\bar{x} \pm z_{\alpha/2} s_{\bar{x}}$ 来做,对本题有

$$s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{\sqrt{0.56g^2}}{\sqrt{101}} = 0.0744618g$$

并由表 A.5 可得

$$z_{\alpha/2} = z_{0.10/2} = 1.645$$

于是 μ 的近似置信区间为

$$18.2g \pm (1.645 \times 0.0744618g)$$

$$18.2g \pm 0.12g$$

或

$$(18.08g, 18.32g)$$

为得到 σ^2 的精确的95%置信区间,我们将 $n-1=100$, $s^2=0.56g^2$ 及由表 A.7 得到的

$$\chi^2_{\alpha/2,\nu} = \chi^2_{0.10/2,100} = 124.3$$

$$\chi^2_{1-\alpha/2,\nu} = \chi^2_{1-0.10/2,100} = 77.93$$

代入置信限(见15.6节)得

$$\left(\frac{100 \times 0.56g^2}{124.3}, \frac{100 \times 0.56g^2}{77.93}\right) = (0.45g^2, 0.72g^2)$$

15.9 一个灯泡生产商想提高他的三种灯泡的平均寿命(200hr).如今开发了一种新产品,检验了其中400个,得到这些灯泡寿命结果为 $\bar{x}=286.2hr$, $s^2=26.32hr^2$.假定灯泡寿命测量值总体为正态分布,试用(15.12)式给出 σ^2 的近似99%置信区间.

解 ☞　将 $s^2=26.32hr^2$, $n=400$ 及由表 A.7 得到的 $z_{0.01/2}=2.575$ 代入(15.12)式,得到 σ^2 的近似99%置信区间为

$$26.32hr^2 \pm \left(2.575 \times 26.32hr^2 \times \sqrt{\frac{2}{400}}\right)$$

$$26.32hr^2 \pm 4.79hr^2$$

或

$$(21.53hr^2, 31.11hr^2)$$

15.10　重复习题15.9的讨论，为得到误差边缘 $E=4.00\text{hr}^2$ 的近似99%置信区间，需多大的样本？

解 ☞　由(15.14)式及 $z_{\alpha/2}=z_{0.01/2}=2.575$，$s^2=26.32\text{hr}^2$ 和 $E=4.00\text{hr}^2$ 得

$$n\approx\left(\frac{\sqrt{2}\times 2.575\times 26.32\text{hr}^2}{4.00\text{hr}^2}\right)^2=574.16\text{ 或 }575$$

于是若误差边缘由 $E=4.79\text{hr}^2$ 减至 $E=4.00\text{hr}^2$，我们需将样本由400至少增加至575.

标准差的置信区间

15.11　在15.8中，我们得到了总脂肪测量值总体 σ^2 的精确的90%置信区间(0.45g^2，0.72g^2)．问该总体 σ 的精确的90%置信区间是什么？

解 ☞　由(15.15)式，σ 的精确的90%置信区间为

$$(\sqrt{0.45\text{g}^2},\sqrt{0.72\text{g}^2})=(0.67\text{g},0.85\text{g})$$

15.12　在习题15.9中我们得到了灯泡寿命测量值的正态分布总体中方差 σ^2 的近似99%置信区间是(21.53hr^2，31.11hr^2)．试用下列两种方法给出该总体 σ 的近似99%置信区间：(a)对此近似值取算术平方根；(b)用(15.16)式计算．

解 ☞　(a) $(\sqrt{21.53\text{hr}^2},\sqrt{31.11\text{hr}^2})=(4.64\text{hr},5.58\text{hr})$

(b) 由(15.16)式及

$$s=\sqrt{26.32\text{hr}^2}=5.13030\text{hr}$$

$$z_{\alpha/2}=z_{0.01/2}=2.575$$

$$2n=2\times 400=800$$

得到近似99%置信区间为

$$5.13\text{hr}\pm\left(2.575\times\frac{5.13030\text{hr}}{\sqrt{800}}\right)$$

$$5.13\text{hr}\pm 0.47\text{hr}$$

或

$$(4.66\text{hr},5.60\text{hr})$$

二项总体比率的置信区间

15.13　例15.13给出快速生长的洋葱种子的发芽比率 p，若想得到 $E=0.02$ 或更小的近似95%置信区间，至少需检验1,801颗种子．对1,801颗种子进行发芽试验，发现有1,423颗发芽，就此结果，求 p 的近似95%置信区间，并说明此区间是否满足最大容许误差边缘 $E=0.02$？

解 ☞　该研究给出 p 的点估计值为

$$\bar{p}=\frac{1{,}423}{1{,}801}=0.790117$$

样本大小满足近似正态方法所需最严格条件(见15.15和15.16节)：

$$(n\bar{p}=1{,}423.0)>15,[n(\bar{q}=1-\bar{p})=378.0]>15$$

及

$$(n=1{,}801)>100$$

由于 n“足够大”，所以可以用点估计法解．将 $n=1{,}801$，$z_{0.05/2}=1.96$ 及 $\bar{p}=0.790117$ 代入(15.19)式，得到 p 的近似95%置信区间为

$$0.790117\pm\left[1.96\times\sqrt{\frac{0.790117(1-0.790117)}{1{,}801}}\right]$$

$$0.79 \pm 0.019$$

因为 0.019 小于 0.20，所以 1,801 的样本大小可以满足 $E=0.02$ 的误差边缘.

15.14 一汽车制造商接到一批 3,000 件新型电子元件(见 11.11 节). 该公司想知道有缺陷元件比率 p 的近似 90%置信区间，误差边缘不超过 $E=0.04$. 如果将第一批元件视为总体，应从这批中抽取一个多大的样本进行检验?

解 ☞ 由于没有 p 的先验点估计值，所以应用保守法解. 将 $z_{0.10/2}=1.645$ 和 $E=0.04$ 代入(15.28)式，得

$$n \approx \left(\frac{1.645}{2\times 0.04}\right)^2 = 422.816 \text{ 或 } 423$$

由于此总体不是很大，根据 15.20 节中的说明，由(15.29)式得到修正的样本量为

$$\text{修正样本量} \approx \frac{423}{1+\frac{423}{3,000}} = 370.727 \text{ 或 } 371$$

15.15 由习题 15.14 的讨论知样本容量为 371 件，其中有 18 件有缺陷. 求总体(批)中有缺陷元件的比率 p 的近似 95%置信区间，并指出样本容量和置信区间能否满足误差边缘 $E\leqslant 0.04$?

解 ☞ 因为讨论中给出了 p 的点估计值：

$$\bar{p} = \frac{18}{371} = 0.0485175$$

样本容量满足近似正态方法的最严格条件(见习题 15.13)：

$$(n\bar{p}=18.0) > 15,\ [n(\bar{q}=1-\bar{p})=353.0] > 15$$

及

$$(n=371) > 100$$

由于 n"足够大"，所以用点估计解法(见 15.15 节)，但因为$(n/N=371/3,000=0.123677)>0.05$，故须用(15.21)式. 将 $n=371$，$N=3,000$，$z_{0.10/2}=1.645$ 及 $\bar{p}=0.0485175$ 代入式中得

$$0.0485175 \pm \left[1.645\times\sqrt{\frac{0.0485175(1-0.0485175)}{371}}\times\sqrt{\frac{3,000-371}{3,000-1}}\right]$$

$$0.049 \pm 0.0172$$

误差边缘 $E\approx 0.0172$ 小于最大可接受的 $E=0.04$.

15.16 在习题 15.15 中，我们知道用大小为 371 的样本给出的误差边缘为 $E\approx 0.0172$，小于最大可接受的 $E=0.04$. 这些结果表明，我们可以用较小的样本来重新讨论. 对同样的置信水平(90%)，试用习题 15.15 中的 $\bar{p}=0.0485175$ 给出满足 $E=0.04$ 的最小样本容量 n?

解 ☞ 由(15.28)式，用点估计解法求 n，并将特定值 $z_{\alpha/2}=z_{0.01/2}=1.645$，$E=0.04$ 和 $p=0.0485175$代入，则得到所需样本容量为

$$n \approx \left[\frac{1.645\sqrt{0.0485175(1-0.0485175)}}{0.04}\right]^2 = 78.07$$

因此，为得到所需的估计性质至少需要 79 件. 这能够满足 15.16 节用点估计解法的保守的要求：$n\geqslant 30$. 但是若想满足更保守的需求：$np\geqslant 10$，n 至少为 206.

二项分布百分数及总数的置信区间

15.17 在习题 15.15 中，我们得到了总体(批)中有缺陷元件比率 p 的近似 90%置信区间：0.049 ± 0.0172. 应用这个样本的结果，求出总体中有缺陷元件百分比 $p\times 100\%$ 的近似 92%置信区间.

解 ☞ 首先求 p 的近似 92%置信区间，仍用有限总体修正因子的点估计解法来计算置信区间

[(15.21)式]. 由表 A.5 知，面积最接近[0.5 − ($\alpha/2 = 0.08/2 = 0.04$) = 0.46]的 z 值为 $z_{\alpha/2} = z_{0.08/2} = 1.75$. 将该值及 $n = 371$, $N = 3,000$, $\bar{p} = 0.0485175$(见习题 15.15)代入式中，得

$$0.0485175 \pm \left[1.75 \times \sqrt{\frac{0.0485175(1-0.0485175)}{371}} \times \sqrt{\frac{3,000-371}{3,000-1}}\right]$$

$$0.049 \pm 0.0183$$

区间端点同乘以 100%(见 15.21 节)，即得到总体百分比 $p \times 100\%$ 的近似 92%置信区间：

$$(0.049 \times 100\%) \pm (0.0183 \times 100\%)$$

$$4.9\% \pm 1.83\%$$

15.18　一名新的候选人正进行总统竞选. 他让民意测验组织进行全国性的随机数字电话拨号调查(见上册习题 3.25)得到潜在的赞成他而反对在职总统的投票人百分比. 为得到潜在投票者总体百分比($p \times 100\%$)的近似 95%置信区间，且误差边缘 $E = 3.0\%$，这个机构需调查多少美国成年人?

解 ☞　由于没有可用的先验点估计值 $\bar{p}$，所以民意测验组织应将 $z_{\alpha/2} = z_{0.05/2} = 1.96$ 和 $E = 0.03$代入(15.28)式，得到

$$n \approx \left(\frac{1.96}{2 \times 0.03}\right)^2 = 1,067.11 \text{ 或 } 1,068$$

15.19　正如习题 15.18 中所建议的，民意测验组织调查了 1,068 名投票者，并知 384 人赞成这位新候选人，545 人赞成现任总统，且有 139 人弃权. 由此结果，给出潜在的投票者总体中赞成候选人的百分比为($p \times 100\%$)的近似 95%置信区间.

解 ☞　首先求出这些投票人的比率 p 的近似 95%置信区间，调查结果表明点估计值为

$$\bar{p} = \frac{384}{1,068} = 0.359551$$

样本 n 满足近似正态方法所需的最严格条件(见 15.16 节)，且由于 n"足够大"，$n \leqslant 0.05N$，故可用(15.19)式. 将 $n = 1,068$, $z_{0.05/2} = 1.96$ 及 $\bar{p} = 0.359551$ 代入式中，得到 p 的近似 95%置信区间：

$$0.359551 \pm \left[1.96 \times \sqrt{\frac{0.359551(1-0.359551)}{1,068}}\right]$$

$$0.36 \pm 0.029$$

在区间端点同乘以 100%，得到 $p \times 100\%$ 的近似 95%置信区间：

$$(0.36 \times 100\%) \pm (0.029 \times 100\%)$$

$$36\% \pm 2.9\%$$

15.20　习题 15.18 和 15.19 中的候选人想知道赞成他而反对现任总统的人数的百分比. 为解决这一问题，下面是习题 15.19 中民意测验调查结果的正确表述吗?

在 1068 各美国成年人的电话民意调查中，有 36%的人表明他们将支持他而反对在职总统. 此一结果在 95%置信水平下有 2.9 个百分数的误差边缘. 理论上 95%的置信水平意味着在 20 次民意调查中将有 19 次得此结果，与实际的全国投票得票率 $p \times 100\%$有不超过(+)或(−)误差边缘的差别.

解 ☞　除了"有 2.9 个百分数的误差边缘"一句外，表述是基本正确的. 误差边缘应为2.9个百分点. 百分点是百分尺度的单位，误差边缘应是按此尺度从点估计 $\bar{p} = 36\%$的两侧2.9个这种单位的基本单位. "2.9 个百分数"的误差边缘意味着由 $\bar{p}$ 到两侧有($0.029 \times 36 = 1.04$)个按此尺度的百分点的尺度的距离.

15.21　在全国民意调查中，为什么大多数民意调查组织认为大小为 1,500 的样本是"足够大"的呢?

解 ☞ 大多数民意测验组织认为以下是实际中的标准目标:即得到一个总体百分比 $p\times100\%$ 的近似95%置信区间最多只能有三个百分点的误差边缘.由习题15.18和15.19结论可知,不考虑被抽样总体的大小,这一目标一般可在样本大小为1500下实现.

15.22 一城镇议会想就本镇"是否应购买公园用地"询问本镇2,000名登记的投票者的一个随机样本.如果议会想得到已登记投票者中同意购买的百分比 $p\times100\%$ 的近似95%置信区间,且误差边缘最多是三个百分点,则需调查的样本为多大?

解 ☞ 因为没有 p 的先验点估计值,所以应用(15.28)式,将 $z_{0.05/2}=1.96, E=0.03$ 代入式中得

$$n\approx\left(\frac{1.96}{2\times0.03}\right)^2=1,067.11 \text{ 或 } 1,068$$

由于此题不是一个大总体的民意测验,可将 $n=1,068$ 和 $N=2,000$ 代入(15.29)式以减小样本量,得到修正样本量:

$$\text{修正样本量}\approx\frac{1,068}{1+\frac{1,068}{2,000}}=696.22 \text{ 或 } 697$$

15.23 习题15.22中的民意测验已进行完毕,并在抽样的697名投票者中有382人回答"同意"购买土地.由此结果,给出本镇登记投票者同意购买的百分数 $p\times100\%$ 的近似95%的置信区间.

解 ☞ 首先求出总体比率 p 的近似95%置信区间,民意测验给出了 p 的点估计值:

$$\bar{p}=\frac{382}{697}=0.5488063$$

样本大小满足近似正态方法条件(见15.16节),但是由于($n/N=697/2,000=0.35$)>0.05,所以必须用(15.21)式,将 $z_{0.05/2}=1.96$、$n=697, N=2,000$ 及 $\bar{p}=0.548063$ 代入式中得

$$0.548063\pm\left[1.96\times\sqrt{\frac{0.548063(1-0.548063)}{697}}\times\sqrt{\frac{2,000-697}{2,000-1}}\right]$$

$$0.55\pm0.0298$$

将区间端点同乘以100%(见15.21节)得 $p\times100\%$ 的近似95%置信区间:

$$(0.55\times100\%)\pm(0.0298\times100\%)$$

$$55\%\pm2.98\%$$

根据镇议会所需(见习题15.22),近似95%区间有"至多为三个百分点"的误差边缘.

15.24 对习题15.14中3,000件的总体(批),总体中有缺陷元件总数的近似92%置信区间为多少?

解 ☞ 习题15.17已给出总体中失效比率 p 的近似92%置信区间 0.049 ± 0.0183.将区间端点同乘以总量 $N=3,000$,即可得到失效元件总数的近似92%置信区间:

$$(0.049\times3,000)\pm(0.0183\times3,000)$$

$$147\pm54.9$$

15.25 就习题15.22和15.23中的2,000名已登记投票者总体,给出同意购买用地的已登记投票者总数的近似95%置信区间.

解 ☞ 用习题15.24中的同样方法,将习题15.23中 p 的近似95%区间元素同乘以总体容量 $N=2,000$,得到这个近似的95%置信区间为

$$(0.55\times2,000)\pm(0.0298\times2,000)$$

$$1,100\pm59.6$$

用捕获-再捕获方法来估计总体大小

15.26　在使用某杀虫剂之前，一种植小麦的农民想估计一下地里的蝗虫数目.在蝗虫的活动范围内随机选择了一块足够小的地方，在此地设置了五个长捕虫网.当他倒净网后，发现捉到了 164 只蝗虫，分别用瓷漆颜料做好标记，然后释放.第二天他又重复此过程，捉到 183 只，其中有 13 只来自有标记的第一样本.试问该块地中蝗虫总数约为多少？

解 ☞　将 $n_1=164, n_2=183, r=13$ 代入(15.30)式得

$$N \approx \frac{(164)(183)}{13} = 2{,}308.6 \text{ 或 } 2{,}309$$

15.27　城市公园管理人员想知道平常早晨公园里学龄前儿童的数目.某一早晨，在四个随机选择的 15 分钟间隔里，他在入口处记下所有不同孩子的名字，共 26 人.用同一方法，下一平常早晨记录下第二个样本为 32 人，其中有 17 人属于第一样本.早晨来公园的学龄前儿童大约有多少人？

解 ☞　将 $n_1=26, n_2=32, r=17$ 代入(15.30)式得

$$N \approx \frac{(26)(32)}{17} = 48.9 \text{ 或 } 49$$

补充习题

附录表 A.7：χ^2 分布的临界值

15.28　利用附录表 A.7 求 $\chi^2_{1-\alpha/2,\nu} = \chi^2_{1-0.100/2,50}$.

答案：34.76

15.29　利用附录表 A.7 求 $\chi^2_{\alpha,\nu} = \chi^2_{0.500,50}$.

答案：20.34

15.30　利用附录表 A.7 求 $\chi^2_{1-\alpha,\nu} = \chi^2_{1-0.500,40}$.

答案：39.34

方差的置信区间

15.31　一昆虫学家对甲虫大小很感兴趣.他取了 20 只甲虫的随机样本，测量它们的翅膀长度(以毫米 mm 为单位)，得到 $\bar{x}=32.4$mm，$s=4.02$mm.假定翅膀长度的总体是正态分布的，总体方差 σ^2 的 95% 置信区间是什么？

答案：(9.35mm，34.46mm)

15.32　若在习题 15.31 所描述的总体中另取一更大的样本($n=40$)，此时的 $s=3.90$mm，试用近似正态方法(见 15.9 节)求总体方差 σ^2 的近似 95% 置信区间.

答案：(10.29mm，25.50mm)

15.33　进一步，在习题 15.31 所描述的总体中取更大样本($n=200$)，此时 $s=3.48$mm，试由(15.12)式给出总体方差 σ^2 的近似 95% 置信区间.

答案：(9.74mm，14.48mm)

标准差的置信区间

15.34　已知雄地鼠的体重是正态分布的.取 12 只雄地鼠的随机样本，得到以克为单位的下列重量：380，385，400，375，420，415，412，401，405，411，410，383.给出总体标准差 σ 的精确的 95% 置信区间([(15.15)式].

答案：(10.79g，25.86g)

15.35　对习题 15.34 中地鼠的体重，假设已取得一较大样本($n=120$)，且已知 $\bar{x}=400.0$g，$s=14.0$g.利用样本标准差的抽样分布[(15.16)式]给出总体标准差 σ 的近似 95% 置信区间.

答案：(12.23g，15.77g)

二项总体比率的近似置信区间

15.36　已知一湖中约有 2,000 条蓝鳃鱼，有放回地从中取一随机样本($n=120$)，该样本中雄性比率为 $\bar{p}=0.45$.试用近似正态方法给出未知总体比率 p，由：(a)点估计解法，(b)保守解法下的近似 95% 置信区

间.

答案:(a)0.45 ± 0.0890, (b)0.45 ± 0.0895

15.37 在习题 15.36 的约为 2,000 条蓝鳃鱼的总体中,假定样本($n = 120$,雄性的 $\bar{p} = 0.45$)是无放回的,试用近似正态方法给出未知总体比例 p,由:(a)点估计解法,(b)保守解法下的近似 95%置信区间.

答案:(a) 0.45 ± 0.0863, (b) 0.45 ± 0.0868

15.38 在例 15.11(a)中,由点估计解法及 $\bar{p} = 0.546667$ 得到 p 的近似置信区间 0.55 ± 0.0797.若我们仍然用点估计解法,那么要想得到 $E = 0.03$ 的近似 95%置信区间需多大的样本容量 n?

答案:1,058

15.39 在例 15.11(b)中,由保守解法及 $\bar{p} = 0.546667$ 得到 p 的近似置信区间 0.55 ± 0.800.若我们仍然用保守解法研究,那么要想得到 $E = 0.03$ 的近似 95%置信区间需多大的样本容量 n?

答案:1,068

二项分布百分数和总数的置信区间

15.40 一汽车销售商想知道在其销售区的 320 名十七岁人中拥有自己的汽车的百分比.他无放回的选取了 50 人的随机样本,发现其中 18 人有自己的汽车.该二项总体百分数的近似 90%置信区间为多少?

答案:$36\% \pm 10.3\%$

15.41 一生理研究员想知道跑步对减肥的效果.他从 18,000 位男性跑步者中随机选取了 80 名为样本,并将其分为两组:一组为一周跑步超过 50 英里,另一组则少于 50 英里.在此样本中他发现有 16 人一周跑步超过 50 英里.求每周超过 50 英里跑步者的百分数 $p \times 100\%$ 的近似 95%置信区间.

答案:$20\% \pm 8.8\%$

15.42 就习题 15.40 中的研究,给出拥有汽车总人数的近似 90%置信区间.

答案:115.2 ± 32.96

15.43 就习题 15.41 中的研究,给出每周跑步超过 50 英里总人数的近似 95%置信区间.

答案:$3,600 \pm 1,548$

用捕获-再捕获方法估计总体大小

15.44 冬天一大群红翅乌鸫聚集在中西部的麦田里,一麦农想知道栖息在他的小林地树上该类鸟的数目.于是他设置了诱饵网,任意放于小林地中,抓到 28 只鸟并做了记号后释放.四天后他以同样的方法抓到 22 只,其中 6 只有标记.问他的小林地中鸟群大约有多少?

答案:103 只

15.45 一位关心青蛙减少的自然资源保护论者想检查一下池塘中的青蛙总体,他的计划中的一部分是每年估计池塘中蝌蚪的数量.任意在池塘布网,捕获了 120 只蝌蚪,并用防水颜料在其身上画了一个小点作为标记后释放.二天后他取第二个样本,抓到 99 只,其中 18 只有标记.池塘中的蝌蚪大约有多少?

答案:660 只

第十六章　单样本的假设检验

16.1　统计假设检验

推断性统计有四个理论部分，其中的三个：概率理论，抽样理论和估计理论前面已经介绍．本章介绍最后一部分：假设检验理论，目的便是检验此理论如何用于**总体参数**：均值 μ，方差 σ^2，标准差 σ 以及比率 p 的**单样本假设检验**．在第十七章将介绍解决估计和假设检验问题的两样本方法，第 18 章将给出解决这两类问题的多样本方法．

虽然估计和假设检验都用经验样本对未知总体特征进行统计推断，但推断的形式完全不同．考虑未知总体参数 θ，在统计估计问题中，样本信息用于估计 θ 的值，并以一个已知的"置信"度代替 θ，此值落在某个区间的边界内（见第 14、15 章）．相比之下，在统计假设检验问题中，关于 θ 的对立的、备择的假设形成一对**统计假设**，样本信息用于决策，并以一个已知"概率"度，接受对立的假设之一，因而它正确．

16.2　零假设和对立假设

在上册的 3.6 节，我们定义的**统计假设**为：关于一个或多个测量总体的未知性质的假定（或猜想），一般是关于总体参数或关于总体由最小值到最大值是如何散布（分布）的．

现在，在研究的更高阶段，我们可以说这一假定并不是关于测量总体的，而是关于描述总体的数学模型的随机变量（见上册 10.1 和 10.2 节）和其概率分布（见上册 10.3 和 10.4 节）．本章讨论模型中的未知参数 θ．

每个假设检验问题都有一对竞争的统计假设：**零假设**和**对立假设**．零假设，记作 H_0，一般是一明确的语句：未知的总体参数 θ 等于某个特殊常数值．由符号表示为：

$$H_0 : \theta = \theta_0$$

因为 θ 和 θ_0 间不存在差别，所以零假设也称为**无差别假设**或**零差别假设**．

对立假设，记为 H_1（或 H_a，H_A），它是关于 θ 的不同于 H_0 的假设．对立假设有四种常用形式：

(1) $H_1 : \theta \neq \theta_0$，也称为**双侧对立**．由于它在 θ_0 两侧讨论与 θ_0 的可能不同，所以这一假设也称为**无方向对立假设**．

(2) $H_1 : \theta < \theta_0$，也称为**单侧对立**．因为它只在 θ_0 一侧讨论与 θ_0 的可能不同，由于所关心的是 θ_0 左侧或小于 θ_0 的值，所以这一假设也称为**左向对立假设**或**小于方向对立假设**．

(3) $H_1 : \theta > \theta_0$，也称为**单侧对立**．因为它只关心 θ_0 的右侧或大于 θ_0 的值，所以这一假设也称为**右向对立假设**或**大于方向对立假设**．

(4) $H_1 : \theta = \theta_1$．在假设检验的研究中，我们不用这种对立假设，但它在确定**假设检验功效**时具有重要地位（见 16.8 节）．

有些统计假设是简单假设，而有些是复合假设．**简单统计假设**明确的说明参数 θ 等于某一特定常数．零假设（$H_0 : \theta = \theta_0$）与以上所列第四个对立假设（$H_1 : \theta = \theta_1$）是简单假设．**复合统计假设**则是简单假设的非特定数的组合．以上所列假设中，$H_1 : \theta \neq \theta_0$，$H_1 : \theta > \theta_0$ 和 $H_1 : \theta < \theta_0$ 是复合假设．因此，举例来说，$H_1 : \theta \neq \theta_0$ 叙述为 θ 不等于 θ_0，其意义是它等于某个别的非特定值．

16.3 零假设的检验

在单样本假设检验之前,就有足够的有关未知总体参数 θ 的信息,使得能够从下列检验统计假设中选出一个:

(1) $H_0:\theta=\theta_0$, $H_1:\theta\neq\theta_0$;

(2) $H_0:\theta=\theta_0$, $H_1:\theta<\theta_0$;

(3) $H_0:\theta=\theta_0$, $H_1:\theta>\theta_0$.

每对假设的选择依赖于因子的种类(见 16.11 节).

在统计假设检验推理中,检验总是从选择假设对开始的,实际上是零假设 H_0 的检验.从感兴趣的总体中选取大小为 n 的随机样本,由样本计算未知参数 θ 的点估计值 θ^*.然后用一个称为**假设检验**的过程比较 θ^* 与零假设中的 θ_0 值.比较过程因选择的对立假设 H_1 变化,但不论使用哪一过程都将得到下列决策之一:

(1) 如果 θ^* 是那样的接近于 θ_0,以至检验规则判断其与 θ_0 是**"一致的"**,则**接受零假设** H_0,因此**拒绝对立假设** H_1.

(2) 如果 θ^* 和 θ_0 是那样的不同,以至检验规则判断其与 θ_0**"不一致的"**,则**拒绝零假设** H_0,因此**接受对立假设** H_1.

假设检验逻辑是基于以数学形式进行的**演绎推理**的推论(见上册 3.7 节和习题 3.6).此过程以零假设实际上作为其起始前提,然后检查此前提的逻辑推论—如果零假设为真,推论也一定为真.特别地,其确定了**条件概率**(见上册 9.2 节):如果已知 $H_0:\theta=\theta_0$ 为真,那么与零假设中的值(θ_0)差别,至少与样本计算得到的(θ^*)值一样的点估计值的概率为多少?或者以符号形式记为:

$$P(\text{与 } \theta_0 \text{ 至少相差 } \theta_0 \text{ 的点估计值} \mid H_0 \text{ 为真}) \tag{16.1}$$

这一条件概率是由具有下列性质的**检验统计量**决定的(见 14.1 节):(1)它考虑到样本点估计值(θ^*)和零假设值(θ_0)的比较;(2)它与假定 H_0 为真时的一个已知的概率分布有关.一般第 14、15 章中所用的样本统计量都可用于假设检验,但以前是统计量使未知参数 θ 置于一个区间内.现在则是将由 H_0 假设的 θ 的特定值(即 θ_0)置于与点估计值 θ^* 比较的统计量中.

16.4 单侧与双侧假设检验

一个假设检验可能是双侧的或是单侧的,这依赖于与零假设 $H_0:\theta=\theta_0$ 配合的对立假设 H_1 的情况.在**双侧假设检验**中,选择无方向或双向的对立假设 $H_1:\theta\neq\theta_0$;在**单侧假设检验**中,可选任一方向的单侧对立假设,如果选 $H_1:\theta<\theta_0$,则单侧检验称为**左侧检验**或**小于方向检验**;如果选 $H_1:\theta>\theta_0$,则单侧检验称为**右侧检验**或**大于方向检验**.

16.5 总体均值 μ 的假设检验:标准差 σ 已知的正态分布总体

为理解假设检验中的概念,在由本节开始一直到 16.18 节的讨论中,我们将考虑从一无限大、均值 μ 未知、标准差 σ 已知的正态分布总体随机抽样的熟悉的例子(已知总体标准差的假定在多数情况下是不切实际的,但它使我们能用熟悉的 Z 统计量和其标准正态分布提出一些概念).

例如,若想知道未知的总体均值 μ 是否等于特定值 μ_0,可选择双侧假设检验:零假设为 $H_0:\mu=\mu_0$,双侧对立假设为 $H_1:\mu\neq\mu_0$.假定零假设为真,则可知(见 13.14 节):如果所有容量为 n 的随机样本来自于一无限大的正态分布总体(均值为 μ,标准差为 σ),且对每一样本计算连续随机变量 $\bar{X}$ 的 $\bar{x}$ 值.那么 $\bar{X}$ 有一正态抽样分布,均值为 $\mu_{\bar{x}}=\mu_0$,标准差(标准误差)为

$\sigma_{\bar{x}}=\frac{\sigma_0}{\sqrt{n}}$.进一步还可假定(见13.14节),如果正态分布变量$\bar{X}$经Z变换[(13.12)式]

$$Z=\frac{\bar{X}-\mu_0}{\sigma_{\bar{x}}} \tag{16.2}$$

后为标准化的.则Z的抽样分布将是$\mu_z=0,\sigma_z=1$的标准正态分布.显然Z统计量满足检验统计量的两个标准(见16.3节):(1)以标准误差(被$\sigma_{\bar{x}}$除)为单位表示点估计值($\bar{x}$)与零假设参数(μ_0)的距离;(2)在假定H_0为真下,有已知的概率分布(标准正态分布).

Z统计量可用于度量H_0为真的可能性.如果对一给定样本计算Z统计量的特定值,记作z^*,若$z^*=0$,样本均值$\bar{x}$一定等于μ_0,因而$H_0:\mu=\mu_0$很可能为真.然而,当z^*是一较大数时,在零的正或负向,即$z^*=a$或$z^*=-a$,则$\bar{x}$和μ_0间有一相当的距离,因此$H_0:\mu=\mu_0$不太可能为真.通过计算P值(见16.6节)或用精度法则(见16.10节),得到(16.1)式的一个值,即是对此可能性的量化.

16.6　P值

假设检验的P值是在给定H_0为真条件下,观测到至少和由样本计算值一样大的检验统计量值的条件概率.考虑双边检验,$H_0:\mu=\mu_0,H_1:\mu\neq\mu_0$.在此,由于关心的是在$\mu_0$两边观测到的差别,因此我们必须考虑正、负的极端$z$值:$z^*=a$和$z^*=-a$.所以必须用标准正态分布来计算$P$值,即同时确定两端极值的概率:

$$P\text{值}=P=P[Z\geqslant(z^*=a)\text{或}Z\leqslant(z^*=-a)\mid H_0\text{为真}] \tag{16.3}$$

此P值如图16-1所示,其中标准正态分布的左、右尾阴影面积之和等于P值.由一般加法法

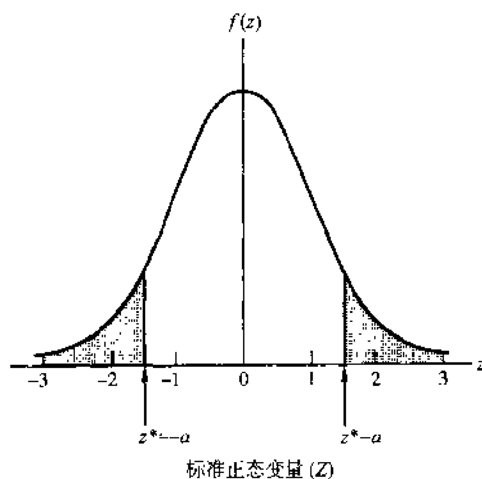


图16-1

则,双侧检验的P值等于下式:

$$P\text{值}=P=P[Z\geqslant(z^*=a)\mid H_0\text{为真}]+P[Z\leqslant(z^*=-a)\mid H_0\text{为真}]$$

由于标准正态是对称的,上式可写成

$$P=2[P(Z\geqslant(z^*=a)\mid H_0\text{为真})] \tag{16.4}$$

考虑假设$H_0:\mu=\mu_0$对$H_1:\mu<\mu_0$的左向检验,仅在观测的z值为负,即$z^*=-a$时要进行检验.如果z^*为正,即样本均值大于假设的均值μ_0,习惯上接受原假设,所以如果z^*为负,即仅考虑负向极端值,则检验的P值为

$$P\text{值} = P = P[Z \leqslant (z^* = -a);\ H_0 \text{为真}] \tag{16.5}$$

这一概率如图 16-2(a)所示,其中标准正态分布的左尾阴影处面积即是 P 值.

而对右向检验假设 $H_0: \mu = \mu_0$ 对 $H_1: \mu > \mu_0$,仅在观测的 z 值为正,即 $z^* = a$ 时要进行检验. 如果 z^* 为负,即样本均值小于假设的均值 μ_0,习惯上接受原假设,所以如果 z^* 为正,即仅考虑正向极端值,则检验的 P 值为

$$P\text{值} = P = P[Z \geqslant (z^* = a)\ \ H_0 \text{为真}] \tag{16.6}$$

这一概率如图 16-2(b)所示,其中标准正态分布的右尾阴影处面积即是 P 值.

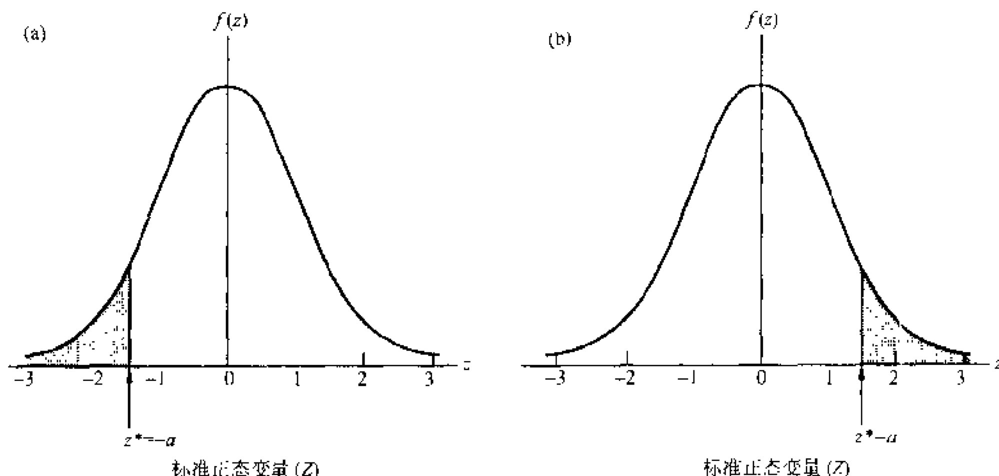


图 16-2

16.7　第 I 类错误和第 II 类错误

一个假设检验作出一个决策:接受 H_0 还是拒绝 H_0. 这一决策是基于样本提供的不完全信息对未知的总体特征的推断,因此总会存在不正确决策的风险. 这种检验可能有四种可能结果,分别为(列在表 16.1 中):

(1) 错误决策:H_0 为真但被拒绝,此类错误称为**第 I 类错误**;

(2) 正确决策:H_0 为不真并被拒绝;

(3) 正确决策:H_0 为真并被接受;

(4) 错误决策:H_0 为不真但被接受,此类错误称为**第 II 类错误**.

表 16.1

决策	总体的实际状态	
	H_0 为真	H_0 不真
拒绝 H_0	第 I 类错误	正确决策
接受 H_0	正确决策	第 II 类错误

16.8　临界值与临界域

P 值(见 16.6 节)可以看作一个描述性统计量,它度量了数据对零假设有多大的支持:P 值越小,支持越小. 但是怎样的支持的水平认为是小的,使得可以拒绝零假设呢?

统计学家们考察决策中错误风险,尤其是第 I 类错误风险:H_0 为真但被拒绝,来回答上面

的问题.已知对一给定样本,P 值越小,H_0 为真的概率也越小.因此犯第Ⅰ类错误的概率也越小,这与另一类错误的风险(概率)作为接受还是拒绝 H_0 的决定因素是一致的.如果 P 小于或等于可接受的最大第Ⅰ类错误风险,用 $\alpha(P\leqslant\alpha)$ 表示,则拒绝 H_0;相反,如果 $P>\alpha$,则认为第Ⅰ类错误的风险太大,于是接受 H_0.最大可接受风险习惯上设为 $\alpha=0.05$,当然如果希望得到一个更保守的决策,将可接受风险设为 $\alpha=0.01$.

为理解 α 的这一用法,考虑图 16-3(a)中给出的例子.由 $H_0:\mu=\mu_0$ 和 $H_1:\mu>\mu_0$ 知,这是一个右向检验.由于 Z 统计量是检验统计量,所以用(16.6)式给出的 P 值:

$$P=P[Z\geqslant(z^*=a)\mid H_0\text{ 为真}]$$

如果该 P 值小于或等于预先给定的 $\alpha(P\leqslant\alpha)$,则拒绝 H_0 而接受 H_1;如果该值大于预先给定的 $\alpha(P>\alpha)$,则接受 H_0.正如以前的 α 定义(见 12.10, 12.13, 14.19 和 15.3 节),α 是右图 16-3(a)给出的标准正态分布右尾区域—z_α 右侧 z 值的集合的面积.如果对所有正的样本值 z^* 都有 $z^*>z_\alpha$ 则 $P[Z\geqslant(z^*=a)\mid H_0$ 为真$]$将小于或等于 α,于是拒绝 H_0.因而称 z_α 右侧阴影区域为**检验的临界域**(或**检验的拒绝域**),称 z_α 为**检验的临界值**.反之,对所有正、负 z^* 值,如果 $z^*\leqslant z_\alpha$,则 $P[Z\geqslant(z^*=a)\mid H_0$ 为真$]$将大于 α,由此接受 H_0.于是称 z_α 左侧的非阴影区域为**检验接受域**(回顾 16.6 节,习惯上在右向检验中如果 z^* 为负,则自然接受 H_0).由图 1 可见一旦设定了 α,由于其代表了当 H_0 为真时**偶然**落入临界域的 z^* 值的相关频数,因而 α 为检验中犯第Ⅰ类错误的概率.

图 16-3(b)给出检验统计量 Z 的左向检验的条件.在此用(16.5)式定义的 P 值来检验 H_0:

$$P=P[Z\leqslant(z^*=-a)\mid H_0\text{ 为真}]$$

同样当 $P\leqslant\alpha$ 时,拒绝 H_0;当 $P>\alpha$ 时,接受 H_0.现在 α 是 $-z_\alpha$ 左侧标准正态分布中的 z 值的集合的面积.于是对所有负的 z^* 值,如果 $z^*<-z_\alpha$,则

$$P=P[Z\leqslant(z^*=-a)\mid H_0\text{ 为真}]$$

将小于或等于 α,因此拒绝 H_0.所以图 16-3(b)中 $-z_\alpha$ 左侧阴影区域是检验的临界域,$-z_\alpha$ 为检验的临界值.反之,对所有正或负的 z^* 值,如果 $z^*\geqslant -z_\alpha$,则 $P=P[Z\leqslant(z^*=-a)\mid H_0$ 为真$]$将大于 α,接受 H_0,于是图 16-3(b)中 $-z_\alpha$ 右侧非阴影区域为接受域(回顾 16.6 节,习惯上在左向检验中,如果 z^* 为正,自然接受 H_0).α 同样是犯第Ⅰ类错误的概率.

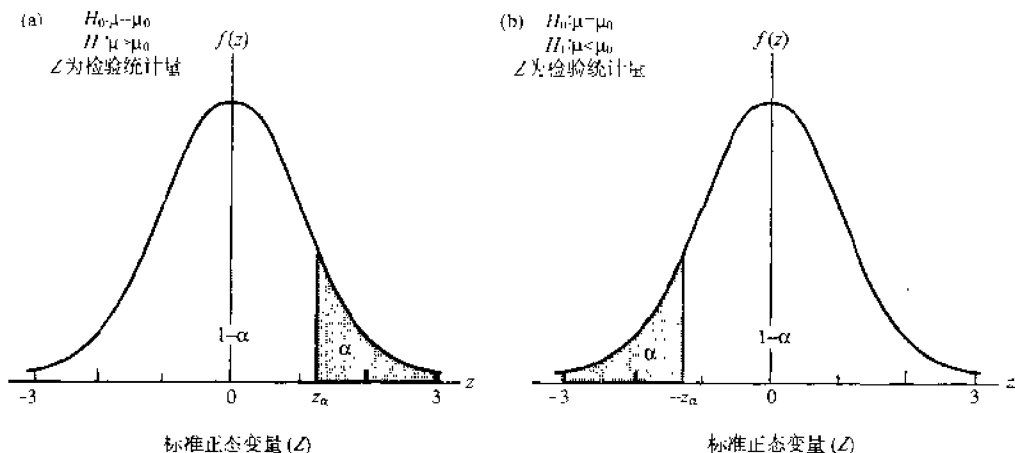


图 16-3

最后，图 16-4 给出检验统计量 Z 的双边检验条件. 在此我们不关心方向[即 z^* 是大于还是小于标准正态分布均值零]，只关心 z^* 是否与均值不同. 因此用(16.4)式中定义的 P 值来检验 H_0：

$$P = 2[P(Z \geqslant (z^* = a) \mid H_0 \text{ 为真})]$$

如前，如果 $P \leqslant \alpha$，拒绝 H_0；如果 $P > \alpha$，接受 H_0. 但是现在 $\alpha = \alpha/2 + \alpha/2$，由于 α 平分到标准正态分布的两个尾部上，即每一尾部为 $\alpha/2$. 对双侧检验问题，α 是分布中 $z_{\alpha/2}$ 右侧和 $-z_{\alpha/2}$ 左侧的 z 值集合的面积(图 16-4 中阴影区域). 如果 $z^* > z_{\alpha/2}$ 或 $z^* < -z_{\alpha/2}$，则 $P = 2[P(Z \geqslant (z^* = a) \mid H_0$ 为真)]小于或等于 α，于是拒绝 H_0. 两个尾部区域之和为检验的临界域，$z_{\alpha/2}$ 和 $-z_{\alpha/2}$ 都为检验的临界值. 反之，如果 z^* 在区间 $-z_{\alpha/2} \leqslant z^* \leqslant z_{\alpha/2}$ 中，则 $P = 2P[(Z \geqslant (z^* = a) \mid H_0$ 为真)]大于 α，于是应接受 H_0. 因此这个区间上的范围(图 16-4 中无阴影区域)为检验的接受域.

单侧假设检验和**双侧假设检验**的说法是指检验统计量的抽样分布中临界域的位置. 在单侧检验中，临界域完全是在分布的一个尾巴(见图 16-3)，因此这样的检验称为单侧检验；在双侧检验中，临界域被平分到分布的两个尾巴(见图 16-4)，因此这样的检验称为双侧检验. 右侧单边检验也称为**右尾检验**或**上尾检验**；左侧单边检验也称为**左尾检验**或**下尾检验**.

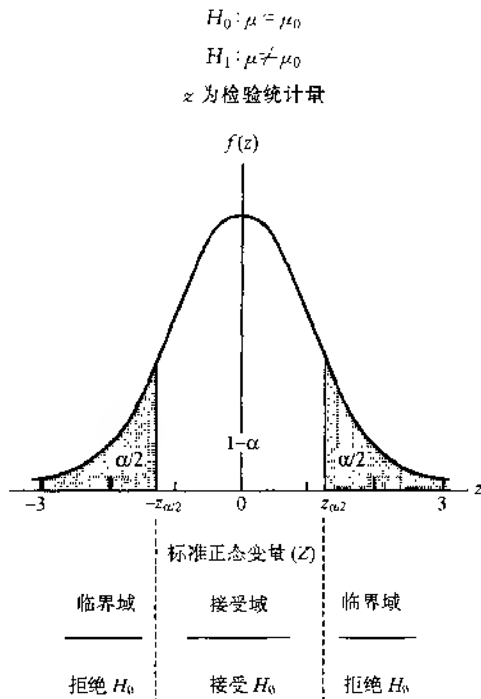


图 16-4

16.9 显著水平

在 16.8 节中我们指出了 α 是第 I 类错误的概率(当零假设为真但被拒绝的概率). 现在我们说 α 也是**假设检验的显著水平**，这是因为它用于评估**样本结果的显著性**，识别与零假设的**显著**(或实际)**不同**. 如果点估计值 $\hat{\theta}^*$ 与假设的参数 θ_0 有很大差别，以至于 $P \leqslant \alpha$，拒绝 H_0，则结果称为**统计显著**；如果 $P > \alpha$ 并接受 H_0，则结果称为**不是统计显著的**. 因此 α 是统计显著与不是统计显著的界限，因此称为**假设检验的显著性水平**或**检验的显著水平**. 在图 16-3 和图 16-4给出的标准正态分布中，临界域即为统计显著区域，而接受域也是非统计显著域.

正如 16.8 节中指明的，α 应在试验前设定为 0.05 或 0.01. 例如当 $\alpha = 0.05$ 时，研究人员应在报告的方法部分说明"统计假设检验是在 0.05 显著水平(或"5% 显著水平")下进行的".

如果 $P\leqslant 0.05$，并拒绝 H_0，则研究人员应说"结果在 0.05 的显著水平下是显著的"，如果 $P>0.05$，并接受 H_0，则研究人员应说"结果在 0.05 的显著水平下不是统计显著的".

"统计显著"和"不是统计显著"仅用于样本结论，而不能用于 H_0、H_1 或它们给出的参数. 因此可以说"样本结论是统计显著的"或"z^* 值是统计显著的"，但不能说"H_0 是统计显著的"或"μ_0 是统计显著的". 然而统计显著的说法确实说明应拒绝 H_0.

由于统计显著程度的存在，因此 $P\leqslant 0.05$ 可说成是"显著的结果"，$P\leqslant 0.01$ 可说成是"高度显著的结果". 在 $\alpha=0.05$ 和 $\alpha=0.01$ 作为显著与非显著间的界限的同时. 研究人员通常指出数据的最高显著水平：$P<0.05$、$P<0.02$、$P<0.01$、$P<0.001$ 等. 这就是为什么尽可能报告数据的实际 P 值，因为它是拒绝 H_0 的最高显著水平 P(即最小 P). 因此，称 P 值为**观测的显著水平或描述性显著水平或达到的显著水平**. 给出实际 P 值，可使读者就结果的显著性作出自己的判断.

16.10　统计假设检验的决策规则

在 16.8 节中，我们介绍了假设检验的基本决策规则：如果 $P\leqslant\alpha$，则拒绝 H_0；如果 $P>\alpha$，则接受 H_0. 一旦在检验前选定了假设检验的组成元素：H_0、H_1、检验统计量和显著水平 α，一般法则便自动成为检验的特定决策规则，这种具体的规则可以表示为二种等价形式：P 值及它们与检验的显著水平的关系，或检验统计量及它们与临界值间的关系.

为理解决策规则，我们首先将 Z 统计量用于图 16-3(a)中的右侧检验. 对给定的水平 α，由表 A.5 得临界 z 值. 因此如果选 $\alpha=0.05$ 作为检验的显著水平，那么两个等价的决策规则为

$$\text{拒绝 } H_0, \quad \text{若 } P\leqslant 0.05$$

或

$$\text{拒绝 } H_0, \quad \text{若 } z^* > (z_\alpha = z_{0.05} = 1.645)$$

两者之所以等价是因为，如果 $z^*>1.645$ 成立，则 $P\leqslant 0.05$(见 16.8 节). 当取 $\alpha=0.01$ 作为显著水平时，两个等价的决策规则为：

$$\text{拒绝 } H_0, \quad \text{若 } P\leqslant 0.01$$

或

$$\text{拒绝 } H_0, \quad \text{若 } z^* > (z_\alpha = z_{0.01} = 2.33)$$

对于图 16-3(b)所示的右尾检验，当选 $\alpha=0.05$ 时，两个等价的决策规则为

$$\text{拒绝 } H_0, \quad \text{若 } P\leqslant 0.05$$

或

$$\text{拒绝 } H_0, \quad \text{若 } z^* < (-z_\alpha = -z_{0.05} = -1.645)$$

当 $\alpha=0.01$ 时，两个等价的决策规则为

$$\text{拒绝 } H_0, \quad \text{若 } P\leqslant 0.01$$

或

$$\text{拒绝 } H_0, \quad \text{若 } z^* < (-z_\alpha = -z_{0.01} = -2.33)$$

对于图 16-4 所示的双边检验，当 $\alpha=0.05$ 时，两个等价的决策规则为

$$\text{拒绝 } H_0, \quad \text{若 } P\leqslant 0.05$$

或

拒绝 H_0,若 $z^* > (z_{\alpha/2} = z_{0.05/2} = 1.96)$ 或若 $z^* < (-z_{\alpha/2} = -z_{0.05/2} = -1.96)$

当 $\alpha = 0.01$ 时,两个等价的决策规则为

$$拒绝\ H_0,\quad 若\ P \leqslant 0.01$$

或

拒绝 H_0,若 $z^* > (z_{\alpha/2} = z_{0.01/2} = 2.575)$ 或若 $z^* < (-z_{\alpha/2} = -z_{0.01/2} = -2.575)$

16.11　统计假设的选择

因为单样本假设检验方法可用于探测参数的变化,所以也称为**检测方法**.在研究的假设检验阶段(见上册 3.8 节),可用来以检测是否存在以参数增加或减少的假设.比如发现一种新型汽油添加剂是否能增加每加仑的英里数.也可在研究探索性阶段(见上册 3.8 节)用这些方法以探测是否发生了什么变化.比如研究一种新型血压药对体温是否有影响?在工业质量控制中这些方法也可用于参数的常规监察过程.比如用于检查工厂的薯片是否与列在包装袋上的参数(比如每份 6 克饱和脂肪)一致.

在研究假设检验的阶段,研究者经常想找到差别,于是一般将研究假设作为对立假设 H_1.所以研究者希望检验会得出拒绝 H_0 并接受 H_1,故我们希望研究者对第 I 类错误(H_0 为真但被拒绝)谨慎对待.如果一个变化的检测有很重要的结果,则应该采用给出最小上界错误的**最大保守**检验方法,即是 $\alpha = 0.01$ 的双侧检验.这是因为要求最大检验统计量值达到拒绝 H_0.同样的原因,按第 I 类错误的**最小保守**检验是 $\alpha = 0.05$ 的单侧检验.

更保守点讲,双侧检验是所有统计假设检验的标准形式,而我们希望的是比较自由的单侧检验.然而当(1)只有一个方向上的变化是重要的(例如,一种新型的减肥药实际减肥多少)或(2)研究的假设预告了一具体的变化方向(例如,经一种新的治疗癌肿瘤会减小)时用单侧假设更合适.而对探索性研究或质量控制监测,任一方向的变化都需要检查时,单侧假设就不合适了.如果用单侧检验,且拒绝 H_0,则结论是具有方向性的(如"有证据表明体重减少").如果用双侧检验,并拒绝 H_0,则结论不具有方向性(如"有证据表明体重变化").一个显著的双侧结论只能"建议"一个变化的方向.

16.12　第 II 类错误概率

16.9 和 16.10 节说明了在统计假设检验中,第 I 类错误是决定接受还是拒绝零假设的决定因素,第 I 类错误的最大风险 α 即是检验的显著性水平:若 $P \leqslant \alpha$,则拒绝 H_0.然而在任何此种统计决策中,往往存在第 II 类可能错误:第 II 类错误(见 16.7 节).现在我们考虑它对假设检验的影响.

第 II 类错误的概率记为 β(罗马小写字母 *beta*).于是假设检验的四种可能结果(见表 16.1)的概率(如表 16.2 所列)为:

表 16.2

决策	总体的实际状态	
	H_0 为真	H_0 不真
拒绝 H_0	第 I 类错误 概率:α	正确决策 概率:$1-\beta$
接受 H_0	正确决策 概率:$1-\alpha$	第 II 类错误 概率:β

(1) $\alpha = P$(第 I 类错误) $= P$(拒绝 $H_0 \mid H_0$ 为真);

(2) $1-\alpha = P$(正确接受 H_0) $= P$(接受 $H_0 \mid H_0$ 为真);

(3) $\beta = P$(第 II 类错误) $= P$(接受 $H_0 | H_0$ 为假);

(4) $1 - \beta = P$(正确拒绝 H_0) $= P$(拒绝 $H_0 | H_0$ 为假).

虽然 α 和 $1-\alpha$ 是已知的,有研究者在检验前设置,但 β 和 $1-\beta$ 的值,若不知道被研究的总体参数 θ 的实际值,是不能确定的.由于这并不是统计假设推断的情形,α 和 β 间的关系已知是相反的:α 越大,β 越小,反之亦然.这一关系之所以成立,是因为 α 越大,接受域越小(见16.8节),因此接受不真的零假设的概率 β 也越小.这一 α 与 β 间的相反关系是在选 α 时必须考虑的又一因素.研究者必须评价两种错误类型的相对严重性,并且认识到若将 α 设置为第 I 类错误风险尽可能小(如,$\alpha = 0.01$ 使第 I 类错误最小),则将增加另一型错误风险.

16.13　使用方风险和生产方风险

在二项分布研究中,我们在批验收抽样(见 11.11 和 11.12 节)中描述了**使用方风险**和**生产方风险**.并定义为:如果称批的接受人为使用方,则使用方风险是接受有高于使用方可容忍的 P 值(不合格品比率)批的概率.如果称批的供应者为生产方,则生产方风险是实际上符合使用方要求的批被拒绝的概率.

批验收抽样方案(见 11.11 节)可应用假设检验方法.比如若使用方需要一批 1000 个标杆的产品,每件直径为 8 毫米(mm),则当这批产品到货后,需要以显著性水平 α 检验这些假设:$H_0: \mu = 8\text{mm}, H_1: \mu \neq 8\text{mm}$.此处生产方风险是第 I 类错误(拒绝真的零假设)的概率,即为 α.而使用方风险是第 II 类错误(接受不真的零假设)的概率,即为 β.

16.14　为何不能证明零假设

零假设 H_0 是关于总体未知特征的一个猜测.因为此特征是未知的(除非检测整个总体),所以不可能绝对确认 H_0 是正确的.取而代之的是用假设检验的方法,对 H_0 作出概率决策,其中要检查的问题为:如果 H_0 为真,那么取得指定样本的概率为多少?如我们所知,假设检验的结果是接受或拒绝 H_0.如果接受 H_0,仅意味着没有足够的理由拒绝 H_0.由于总是存在未知的第 II 类错误(接受错误的零假设)的概率 β,所以不能证明 H_0 是正确的.类似地,如果拒绝 H_0,因为总存在已知的第 I 类错误(拒绝正确的零假设)的概率,所以也不能证明 H_0 是错误的.

16.15　古典推断与 Bayes 推断

本书介绍的推断性统计是**古典推断**的组成部分.到目前为止,在第十四、十五以及十六章的推断性统计介绍中,我们从古典推断的估计和假设检验两方面,解决了有关总体参数的单样本问题.在这些问题中可以看出,我们总是讨论含有未知参数 θ 的总体,而这个未知参数假定是一个常数.由总体的随机样本得到点估计值 $\hat{\theta}^*$,其为 θ 的估计量——随机变量 $\hat{\theta}$ 的特定值.由样本信息,我们用估计方法以一定的置信度确定了含有 θ 值的 $\hat{\theta}$ 的计算区间,用假设检验方法解决了下列形式的问题:若 θ_0 为未知参数值,则将 $\hat{\theta}^*$ 作为点估计值的概率为多少?估计和假设检验方法都是**客观概率**(见上册 8.7 节)的基础,其中的概率是由古典或相对频数概率函数来确定的.

虽然古典推断是统计推断应用最广的一种形式,但还有另一种重要的统计推断形式,这便是 **Bayes 推断**或 **Bayes 决策分析**或**统计决策分析方法**.这种推断形式的目的是为了在下述情况下作出决策,当存在不同的可能行动,且为要从中作出选择所需的信息不完全或不确定时.在 Bayes 推断中,未知参数视为随机变量,而不是常数.将主观概率(见上册 8.7 节)指派的这些随机变量的可能值,使可确定不同行为过程的潜在收益和损失.因为 Bayes 定理(见上册9.9节)用于将指派的主观概率变换为所需新样本信息概率,所以这种推断形式称为 Bayes 推断.

为理解推断的这两种形式的区别及其在实际中的应用,我们再次用在重黏土中生长得很

好的胡萝卜样本(见习题 12.8).用古典推断,种子公司假定在通常生长条件下,胡萝卜新品种有未知的不变均值 μ 的长度测量值总体.由估计方法,他们可以以一定置信度得到一个含有 μ 的X 区间,用假设检验方法,在显著水平 α 下可以检验关于 μ 的假设:$\mu=\mu_0$.

现在如果我们将胡萝卜例子从总体特征转移到市场中,则 Bayes 推断可派上用场.在简单例子中,可用于确定公司应准备多少袋胡萝卜出售,如果需求量很低但却准备很多袋,则会遭受损失;如果需求量很高但又准备得很少,也会遭受损失.将潜在的销售袋数作为需求量,是参数 θ,并将其看作有不同可能水平(如 10,000,20,000 或 30,000)的随机变量,对每一水平指定一个需求量的猜测的主观概率,由此及已知的生产费用可以用 Bayes 方法对选择的行为(如生产的种子的不同数目)估计潜在的收益和损失.尽管存在不确定信息,但这些估计仍然能作出合理的产量决策.为改进这一信息,可以取市场调查样本,用 Bayes 定理修改主观概率.

本书只讨论古典推断,而 Bayes 推断的介绍可在一般的工商业统计教材中找到.

16.16　检验零假设的步骤

由 Z 统计量解决有关未知总体参数 θ 的单样本假设检验问题步骤如下:

(1) 选择零假设 H_0 和对立假设 H_1(见 16.2 和 16.11 节);

(2) 选择显著水平 α(见 16.9 和 16.11 节);

(3) 决定检验统计量(见 16.3 节),由此统计量及 α 来确定检验的决策规则,并用 P 值或临界值(见 16.10 节)描述;

(4) 从总体取一随机样本,并从样本计算检验统计量的值,若可能,计算 P 值;

(5) 由样本结果和决策规则决定是拒绝还是接受零假设(见 16.10 节).

例 16.1　如果一位效力于电池生产商的化学家接到一个为开发改进计算器电池的问题,要求改进的电池比现有电池使用时间"显著的长".已知现有计算器中,电池寿命的度量是正态分布的,$\mu=100.3$(min),$\sigma=6.25$(min).现在开发了一种改进电池,在理论上可能持续更长时间.由初步检验可以假定其寿命度量也是正态分布,$\sigma=6.25$(min).为检验 $H_0:\mu=100.3$(min),选取了一个 $n=15$ 的改进电池的样本,且得到 $\bar{x}=105.6$(min).试在 $\alpha=0.01$ 下作 H_0 的双侧检验,并用 P 值叙述决策规则.按以上步骤给出解.

解 ☞　(1) $H_0:\mu=100.3$(min), $H_1:\mu\neq100.3$(min);

(2) $\alpha=0.01$;

(3) 因为总体为正态分布且 σ 已知,所以我们用 Z 统计量[(16.2)式].当 $\alpha=0.01$ 时,由 P 值描述的双侧检验决策规则(见 16.10 节)为

$$拒绝 H_0,若 P\leqslant 0.01$$

(4)检验统计量的特定值为

$$z^*=\frac{\bar{x}-\mu_0}{\sigma_{\bar{x}}}=\frac{105.6\text{min}-100.3\text{min}}{6.25\text{min}/\sqrt{15}}=3.284,或 3.28$$

为确定 P 值,利用(16.4)式:

$$P=2[P(Z\geqslant(z^*=a)\mid H_0 为真)]$$

因为 Z 是连续随机变量,该公式等价于

$$P=2[P(Z>(z^*=a)\mid H_0 为真)]$$

由表 A.5 可见,在区间 $0\leqslant z\leqslant3.28$ 上标准正态分布面积(概率)为 0.4995.因此,由于零的右侧总面积为0.5,故

$$P=2[P(Z>3.28\mid H_0 为真)]$$
$$=2(0.5-0.4995)=0.001$$

(5)因为 $P<0.01$,所以拒绝 H_0 而接受 H_1.在 0.01 的显著水平下样本平均寿命和假设平均寿命间存在显著差别,而且结果表明改进的计算器的电池有更长的寿命.

例 16.2　重复例 16.1 的电池研究,并将样本容量增至 $n=20$,得到 $\bar{x}=105.0$min.仍假定改进电池的寿

命服从正态分布,且 $\sigma=6.25\text{min}$.在 $\alpha=0.05$ 下,作 $H_0:\mu=100.3\text{min}$ 的单侧检验,并用 P 值叙述决策规则.按前面列出的步骤给出所得解.

解 ☞　(1) $H_0:\mu=100.3\text{min}$.因为只有寿命增加是重要的,且理论上有理由相信是增加的,因此对立的右尾假设:$H_1:\mu>100.3\text{min}$;

(2) $\alpha=0.05$;

(3) 用 $\alpha=0.05$ 时的 Z 统计量,用临界值表示的右尾决策规则(见 16.10 节):

$$\text{拒绝 } H_0\text{,若 } z^* > 1.645$$

(4) Z 统计量的值为

$$z^* = \frac{\bar{x}-\mu_0}{\sigma_{\bar{x}}} = \frac{105.6\text{min}-100.3\text{min}}{6.25\text{min}/\sqrt{20}} = 3.363\text{,或 } 3.36$$

为得到 P 值,用(16.6)式的等价形式

$$P = P(Z > 3.36 \mid H_0 \text{ 为真})$$

由例 16.1 中方法得到 P 值,由表 A.5 得适当值为

$$P = 0.5 - 0.4996 = 0.0004$$

(5) 因为 $z^*>1.645$,所以拒绝 H_0 接受 H_1.这一结论可由 P 值小于 0.05 予以确立.于是可告知公司在显著水平 0.05 下可以认为新电池寿命更长些.

16.17　用 $\bar{X}$ 作为检验统计量的假设检验

如果随机样本取自正态分布总体,均值 μ 未知,标准差 σ 已知,则可直接用随机变量 $\bar{X}$(而不是变换为 Z)作为检验统计量.为理解这一用法,考虑右尾检验 $H_0:\mu=\mu_0$ 对 $H_1:\mu>\mu_0$.首先由(13.36)式可知

$$P(\bar{X} > \bar{x}_\alpha) = P\left(Z > \frac{\bar{x}_\alpha - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) = P(Z > z_\alpha) = \alpha$$

由此可见

$$z_\alpha = \frac{\bar{x}_\alpha - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

用 μ_0 代 $\mu_{\bar{x}}$ 并解出 $\bar{x}_\alpha$ 为

$$\bar{x}_\alpha = \mu_0 + z_\alpha\sigma_{\bar{x}} \tag{16.7}$$

因此,如果我们得到 $\bar{X}$ 的一个特定样本值,记为 $\bar{x}^*$,则用临界值表示的右尾检验决策规则为

$$\text{拒绝 } H_0\text{,若 } \bar{x}^* > (\bar{x}_\alpha = \mu_0 + z_\alpha\sigma_{\bar{x}})$$

类似地,可给出左尾检验和双侧检验临界值决策规则,略去推导过程,左尾检验规则为

$$\text{拒绝 } H_0\text{, 若 } \bar{x}^* < (\mu_0 - z_\alpha\sigma_{\bar{x}})$$

双侧检验规则为

$$\text{拒绝 } H_0\text{, 若 } \bar{x}^* > (\bar{x}_{\alpha/2} = \mu_0 + z_{\alpha/2}\sigma_{\bar{x}}) \text{ 或 } \bar{x}^* < (\bar{x}_{1-\alpha/2} = \mu_0 - z_{\alpha/2}\sigma_{\bar{x}})$$

例 16.3　用 $\bar{X}$ 为检验统计量,重复例 16.2 的右尾检验,并用 16.16 节中的求解步骤给出解.

解 ☞　(1) $H_0:\mu=100.3\text{min}$, $H_1:\mu>100.3\text{min}$;

(2) $\alpha=0.05$;

(3) 因为总体是正态分布且 σ 已知,于是可用 $\bar{X}$ 作为检验统计量.当 $\alpha=0.05$ 时,由表 A.5 知 $z_\alpha=z_{0.05}=1.645$.用临界值表示的右尾决策规则为

拒绝 H_0，若 $\bar{x}^* > \left[\bar{x}_\alpha = \bar{x}_{0.05} = \mu_0 + 1.645\sigma_{\bar{x}} = 100.3\text{min} + 1.645\left(\frac{6.25\text{min}}{\sqrt{20}}\right) = 102.6\text{min}\right]$

(4) $\bar{X}$ 统计量的值为 $\bar{x}^* = 105.0\text{min}$，$P$ 值同例 16.2：$P = 0.0004$.

(5) 因为 $\bar{x}^* > 102.6\text{min}$，所以在 0.05 显著水平下仍是拒绝 H_0，接受 H_1. 这同样可由 $P < 0.05$ 证实.

16.18 检验的功效，操作特性曲线和功效曲线

称正确拒绝错误零假设的概率 $1-\beta$(见 16.12 节)为**假设检验的功效**. 虽然在总体参数 θ 未知下无法确定实际的检验功效，但通常可以在 θ 的可能取值范围内来确定 β，进而确定 $1-\beta$. 如果用 θ_x 记可能的取值，则用 $\beta(\theta_x)$ 和 $1-\beta(\theta_x)$ 表示 β 和 $1-\beta$ 的计算值.

对给定的 α 和样本容量 n，$\beta(\theta_x)$ 图称为**操作特性曲线**或 **OC 曲线**，而将同一条件下 $1-\beta(\theta_x)$ 图称为**功效曲线**或**功效函数**. 这些曲线是用于检测在 θ_x 可能取值范围内检验对错误零假设的灵敏性，也显示了 α 或 n 的可能变化对 $\beta(\theta_x)$ 和 $1-\beta(\theta_x)$ 的影响. 在习题 16.5 至 16.9 中我们将对总体均值给出此类曲线[$\beta(\mu_x)$ 图和 $1-\beta(\mu_x)$ 图].

例 16.4 电池制造商(见例 16.1,16.2 和 16.3)想改进电池使其在计算器中平均使用时间至少比现在多 5min，或平均值至少为105.3min. 对例 16.2 和 16.3 中的右尾检验，当 $n=20$，$\alpha=0.05$ 时，如果改进的电池寿命总体的均值实际上为105.3min，而不是零假设中的 100.3min 时，则第 II 类错误概率为多少？仍假定总体标准差 $\sigma=6.25\text{min}$，总体均值为 100.3min 或 105.3min.

解 图 16-5 列出了得到均值 $\bar{X}$ 的两种可能抽样分布条件：下面的分布是由 $H_0: \mu=\mu_0=100.3\text{min}$ 定义的，均值为 $\mu_{\bar{x}}=\mu_0=100.3\text{min}$，标准差 $\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}=\frac{6.25\text{min}}{\sqrt{20}}$；上面的分布是由简单对立假设 $H_1: \mu=\mu_1=105.3\text{min}$定义的，均值为 $\mu_{\bar{x}}=\mu_1=105.3\text{min}$，标准差同上. 由例 16.3 可知，当 $\alpha=0.05$ 时，下面的零假设定义的分布的临界值为 $\bar{x}_{0.05}=102.6\text{min}$. 于是可知若总体均值实际上为 105.3min，则第 II 类错误的概率(接受错误零假设)是上面抽样分布中 102.6min 左侧的灰色区域，其中 $\beta(\mu_1)=\beta(105.3)$(见 16.12 节). 该结论之所以成立，是因为灰色区域表示了上面分布的所有可能落在检验的接受域中的 $\bar{x}$ 值.

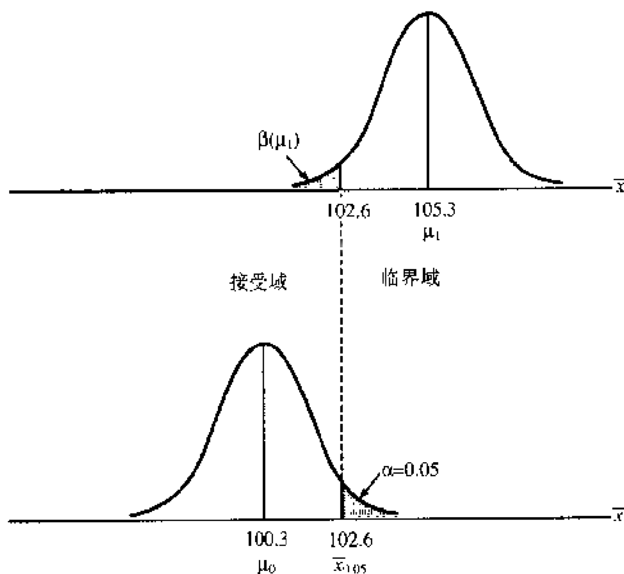


图 16-5

为得到 $\beta(105.3)$值，对上边分布用 Z 变换，得出 z_1(其中下标 1 表示是对立分布的)：

$$z_1 = \frac{102.6\text{min} - 105.3\text{min}}{6.25\text{min}/\sqrt{20}} = -1.932 \text{ 或 } -1.93$$

由表 A.5 得

$$\beta(105.3) = 0.5 - 0.4732 = 0.0268$$

这一结果表明,对于 $\alpha=0.05$ 的右尾检验,如果总体均值实际是 105.3min,则有 2.68% 的可能性导致第 II 类错误.

例 16.5 对于例 16.4 中的右尾检验,若总体均值实际为 $\mu=\mu_1=105.3\text{min}$,试问检验的功效为多少?

解 ☞ 假设检验的功效是 $1-\beta(\theta_r)$,所以此时检验的功效为

$$1-\beta(\mu_1)=1-\beta(105.3)=1-0.0268=0.9732$$

这一结果表明,对于此检验,若 μ 实际上为 105.3min,则 97.32% 的检验正确拒绝错误零假设.

16.19 总体均值 μ 的假设检验:取自未知标准差 σ 的正态分布总体的小样本($n<30$)

到目前为止,在熟悉而不易实现条件:取自 μ 未知、σ 已知的正态分布总体的任何大小样本下,讨论了单样本假设检验.现在讨论更一般情况下的单样本假设检验,即:取自 μ 和 σ 均未知的正态分布总体的小样本($n<30$).在此条件下,虽然所有单样本假设检验的逻辑结构保持不变,但方法的基础需要改变.这是因为,在 σ 未知以及小样本下,已不适合再用 $Z=\frac{\bar{X}-\mu_0}{\sigma_{\bar{x}}}$ 作为统计量了.

现在由于在这些条件下估计的 μ 是真值(见 14.21 节),所以合适的检验统计量是以 μ_0 代替 μ 的 t 统计量[(14.9)式]:

$$T=\frac{\bar{X}-\mu_0}{S_{\bar{x}}}=\frac{\bar{X}-\mu_0}{S/\sqrt{n}} \tag{16.8}$$

可取特定值

$$t=\frac{\bar{x}-\mu_0}{s_{\bar{x}}}=\frac{\bar{x}-\mu_0}{s/\sqrt{n}} \tag{16.9}$$

这一统计量满足检验统计量的两个标准:(1)可以将许样本点估计值($\bar{x}$)和参数的零假设值(μ_0)作比较;(2)在 H_0 为真假定下,t 统计量有已知的概率分布:$\nu=n-1$ 的 t 分布(见 14.15 节).

16.20 t 统计量的 P 值

如果对任一给定容量为 n 的样本,我们将计算的 t 统计量的特定值记为 t^*.因此,用 t 统计量的假设检验 P 值(见 16.6 节)是下面的条件概率:给定 $H_0:\mu=\mu_0$ 为真,在 $\nu=n-1$ 的 t 分布中得到至少与样本计算值一样大的 t^*.下边就三种对立假设 $H_1:\mu\neq\mu_0$,$H_1:\mu<\mu_0$ 和 $H_1:\mu>\mu_0$ 如何计算 P 值.

对于 $H_0:\mu=\mu_0$ 和 $H_1:\mu\neq\mu_0$ 的双边检验,我们必须在适合的 t 分布中找正、负极端值 $t^*=a$ 和 $t^*=-a$,由此 P 值为

$$P=P[(T\geqslant(t^*=a)\text{ 或 }T\leqslant(t^*=-a)\mid H_0\text{ 为真})]$$

并且因为 t 分布是对称的,上式可写作(见 16.6 节)

$$P=2[P(T\geqslant(t^*=a)\mid H_0\text{ 为真}] \tag{16.10}$$

对于 $H_0:\mu=\mu_0$ 和 $H_1:\mu<\mu_0$ 的左尾检验,我们仅关心负向 t^* 极端值,由此 P 值为

$$P=P[T\leqslant(t^*=-a)\mid H_0\text{ 为真}] \tag{16.11}$$

对于 $H_0:\mu=\mu_0$ 和 $H_1:\mu>\mu_0$ 的右尾检验,我们仅关心正向 t^* 极端值,由此 P 值为

$$P = P[T \geqslant (t^* = a) \mid H_0 \text{为真}] \tag{16.12}$$

表A.6(**t分布的临界值**)只限于对给出的t分布以及对给定分布提供t_α和$t_{\alpha/2}$值,因此不可能用它得到准确的P值.但可用表A.6得到近似的P值,或把P值指在一个P值范围内(许多用于计算的统计程序都给出了准确P值).

例16.6　样本($n=20,\bar{x}=4.0,S=0.83$)来自μ和σ都未知的正态分布,若用t统计量对此样本做$H_0:\mu=3.6$右尾检验,用表A.6及(a)插入值法(b)界定法给出观测的t^*的近似P值.

解　(a)由(16.9)式及$\bar{x}=4.0,\mu_0=3.6$和$s_{\bar{x}}=\frac{s}{\sqrt{n}}=\frac{0.83}{\sqrt{20}}=0.1856$得

$$t^* = \frac{4.0-3.6}{0.1856} = 2.1552 \text{ 或 } 2.155$$

由表A.6可知,($\nu=n-1=20-1=19$)的t分布,这个t^*值在$t_{0.025,19}=2.093$和$t_{0.01,19}=2.539$之间,于是P值[(16.12)式]

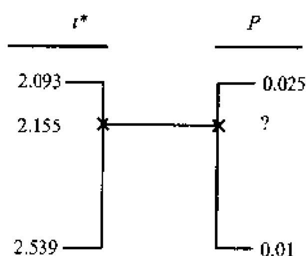


图16-6

$$P(T \geqslant 2.155 \mid H_0 \text{为真})$$

应在0.025和0.01之间.用标准线性插入序列(见例14.15)可得近似值.首先画两条平行尺度:一个是t^*值,另一个是P值(见图16-6).因为$t^*=2.155$是$t^*=2.093$和$t^*=2.539$之间距离的0.0621/0.446,于是相应的P值是$P=0.025$和$P=0.01$之间距离的0.062/0.446,因此

$$P \approx 0.025 - \left[\frac{0.062}{0.446} \times (0.025-0.01=0.015)\right] = 0.0229, \text{ 或 } 0.023$$

(b)由(a)可见P值是在0.025和0.01之间范围中,这可写作

$$0.01 < P < 0.025$$

16.21　t统计量的假设检验决策规则

由于对于t统计量用表A.6仅能给出近似或划界的P值,所以我们仅介绍临界值决策规则.不像Z统计量的决策规则,对任一水平α,给出特定临界值$\pm z_\alpha$或$\pm z_{\alpha/2}$,t统计量的临界值决策规则,对每一样本容量,自由度($\nu=n-1$),给出不同临界值$\pm t_{\alpha,\nu}$或$\pm t_{\alpha/2,\nu}$.所以在此我们给出决策规则的一般形式.

对于右尾检验:

$$\text{拒绝 } H_0, \quad \text{若 } t^* > t_{\alpha,\nu}$$

对于左尾检验:

$$\text{拒绝 } H_0, \quad \text{若 } t^* < -t_{\alpha,\nu}$$

对于双侧检验:

$$\text{拒绝 } H_0, \text{若 } t^* > t_{\alpha/2,\nu} \text{ 或 } t^* < -t_{\alpha/2,\nu}$$

例16.7　就例16.6中取自μ和σ都未知的正态分布总体的样本($n=20,\bar{x}=40,s=0.83$),用$\alpha=0.05$和临界值决策规则,作$H_0:\mu=3.6$的右尾检验,并用16.16节中步骤给出解.

解　(1)$H_0:\mu=3.6$, $H_1:\mu>3.6$;

(2) $\alpha=0.05$;

(3) 因为总体是正态分布且σ未知,所以当$\alpha=0.05,\nu=n-1=20-1=19$时,由表A.6得$t_{0.05,19}$值,用$t$统计量[(16.8)式]得右尾决策规则:拒绝$H_0$,若$t^*>(t_{0.05,19}=1.729)$.

(4) 由例16.6知:$t^*=2.155$.从表A.6可见,对于$\nu=19$的分布,此t^*值在$t_{0.025}$和$t_{0.01}$之间.于是由界定法(见16.20节),P值在$0.01<P<0.025$的范围中.

(5) 因为$t^*>1.729$,所以拒绝H_0接受H_1.这一结果可由$P<0.05$来确认.

16.22 β,$1-\beta$,功效曲线和 OC 曲线

虽然存在用表 A.6 来近似 β 值以及 $1-\beta$ 值的方法,但是对于功效曲线和 OC 曲线需求精确值,故有必要用计算机统计程序产生任意 t 分布的 t 值.

16.23 总体均值 μ 的假设检验:来自任意分布总体的大样本($n \geqslant 30$)

有时候不知道也不能合理假定总体是正态分布的,然而当样本很大时($n \geqslant 30$),则由中心极限定理知 $\bar{X}$ 的抽样分布是近似正态分布(见 13.18 节)[如果样本来自的有限总体是无替换的,则总体容量 N 必须为样本大小 n 的"至少两倍大"(见 13.17 节)].如果这一情况属实,则当 $\bar{X}$ 由 Z 变换[(16.2)式]为标准化后,所得的 Z 分布便近似为标准正态分布(见 13.14 节).检验假设的下一步依赖于总体标准差 σ 是否已知.

当 σ 已知时,可用 16.5 节至 16.7 节所描述的方法检验零假设 $H_0:\mu=\mu_0$.

例 16.8 例 14.9 给出了州高速公路的下坡路段速度总体均值 μ 的近似 95% 置信区间,为 66.3mph ± 1.76mph.此区间是由 $n=85$ 辆车的样本给出 $\bar{x}=66.3$mph 而得的.我们知道总体标准差为 $\sigma=8.3$mph,但不知总体是否为正态分布.现在我们想知道在这一段高速公路上汽车是否比限制速度 65.0mph 显著的快,我们用样本和总体信息在 $\alpha=0.05$ 下,用临界值决策规则对 $H_0:\mu=65.0$mph 作右尾检验,并用 16.16 节中步骤给出解.

解 ☞ (1) $H_0:\mu=65.0$mph, $H_1:\mu>65.0$mph;

(2) $\alpha=0.05$;

(3) 由于 σ 已知,样本足够大($n \geqslant 30$)且 N 为 n 的至少两倍大,所以可应用中心极限定理并假定 X 是近似正态分布.用 Z 检验统计量,当 $n=85$, $z_\alpha=z_{0.05}=1.645$ 时,右尾决策规则用临界值描述为:拒绝 H_0,若 $z^*>1.645$.

(4) Z 统计量的值为

$$z^*=\frac{\bar{x}-\mu_0}{\sigma_{\bar{x}}}=\frac{66.3\text{mph}-65.0\text{mph}}{8.3\text{mph}/\sqrt{85}}=1.444,\ 或\ 1.44$$

由表 A.5 知,近似 P 值为

$$P \approx P[Z \geqslant (z^*=a) \mid H_0\ 为真]=P(Z>1.44)=0.5-0.4251=0.0749$$

(5) 因为 $z^*<1.645$,所以接受 H_0.这一决策可由近似 P 值大于 0.05 予以确认.因此在 0.05 显著水平下,没有明显的证据说明汽车行驶快于限制速度.

当 σ 未知且可假定总体为正态分布时,如果是大样本,则可由下述方法检验零假设:(1)若有精确的 t 值,则用 t 分布的精确解;(2)若此 t 值是近似值,则用 t 分布的近似解;(3)用 Z 分布的近似解.这三种方法已在估计问题的 14.23 节中介绍过.

当 σ 未知且总体明显不是正态分布时,如果样本很大,则可用 Z 分布的近似解,将 $S_{\bar{x}}$ 作为 $\sigma_{\bar{x}}$ 的估计值来检验零假设.此即估计问题的 14.24 节中所描述的方法.

16.24 单样本参数假设检验的假定条件

在此之前,本章描述的假设检验方法是在基于**参数假定**的**参数统计方法**.之所以这样称呼,是因为所检验的统计假设及许多假定涉及到参数.在上册 3.13 节中我们就说过:参数统计方法(或参数统计)是基于有关总体以及所考察的样本特征的非常精确、非常严格的假定,这些假定称为参数假定,描述了所研究的特征例如参数的性质,分布的形状,还须指明抽样的类型.

现在在假设检验讨论中针对此点,我们可以进一步叙述为:这些假定特定数学模型的规定性质是检验的基础.应用数学模型并得到有意义的概率决策,必须满足模型的规定假定.

16.5 节中描述的用 Z 统计量得到精确解的数学模型,已经假定了被研究总体是正态分布,且标准差 σ 已知,也假定了总体是以间隔水平或比例水平度量的(见上册 2.6 和 2.7 节).取任意容量为 n 的简单随机样本(见 13.3 节)满足独立性假定(见 13.4 节),并且对此模型进

一步假定,如果对有限总体进行无替换的抽样,若 $n>0.05N$,则可用有限总体修正因子来计算 $\sigma_{\bar{x}}$(见13.11节).最后,对此及其他统计检验,通常假定每一次都努力用无偏估计量(见14.2节),使测度量偏差(见上册2.13节)和抽样偏差(见上册3.22节)最小.

在16.19节用 t 统计量得到精确解模型中,同样假定了所研究总体为正态分布,但假定 σ 未知,且样本很小($n<30$).上面描述的适用于 Z 统计量模型的其他假定这里同样要求.

在16.23节所描述的用 Z 和 t 统计量来得到精确解和近似解模型,一般假定从总体抽取一个大样本($n\geqslant30$).如果是这样.且假定 σ 已知、总体分布未知,则中心极限定理保证了随机变量 $\bar{X}$ 是近似正态分布的,于是,Z 有近似标准正态分布.当 σ 未知,但假定总体是正态分布时,同用 Z 统计量得出近似解一样,可用 t 统计量得出精确解或近似解.当 σ 未知且总体不能假定为正态分布时,则可用 Z 统计量得出近似解.在前节中模型要求的其他假定在此也同样要求.

16.25 违背假定的情况

一般地,如果我们可以合理的认为,假设检验的假定已经满足,并且若 $P\leqslant\alpha$,知道这一极端的抽样结果不是因为方法有缺陷,而若 H_0 为真一定是因为时有一个很不平常的事件发生,或事实上 H_0 不真.但如果检验的假定中有严重的违背,那么又是什么原因呢?

对于**正态性假定**(假定总体为正态分布),本章适合于此所讨论的检验可视为**稳健的**(见14.15节).这意味着对正态性的中等偏离对结果影响并不严重.由于中心极限定理允许使用近似 Z 解(见16.23节),所以在大样本情况下($n\geqslant30$)可忽略该假定.然而在 σ 未知的小样本情形中,需要精确的 t 解(见16.19节)并且如果正态性偏离严重,则不能用此解,除非有可能**变换**数据(见上册3.13节).

因此,分析这样的小样本数据,第一步便是看样本分布的形状.相对频数分布(见上册4.3节)或茎叶图可以很快给出分布是对称的还是有偏斜、单峰还是多峰.另一方法是在**正态概率纸**上画出样本的累积百分数分布(见上册4.18节):在概率纸上画出的正态分布的累积百分数分布应是一直线.还有一种"拟合优度"检验(见20.3节),比较样本频数与期望的正态分布频数.

如果这些检验表明存在与正态性的严重偏离,则下一步(见上册图3.3)便是试图对数据进行变换.在此变换中,通常用线性测量尺度(见上册2.1节)将数据变换到另一测量尺度,目标就是寻求一个能满足所偏离假定的尺度,这里并不存在变换规则,但一些统计书中针对不同偏离给出了许多方法.例如,如果样本分布为正偏的(见上册习题5.6),则建议用**对数变换**:样本中的每个测量值变换为它的对数值(一般以10为底).或者若数据是离散的测量值(见上册2.8节)且趋向于 Poisson 分布(见11.27节),可取每一样本值的平方根,有近似正态分布.

如果变换成功且基本满足检验的假定,则可对变换后数据进行假设检验,描述结果的报告中必须指出所做的变换.

如果原始数据严重偏离参数假定,且找不到合适变换以满足假定,则不能进行参数检验.于是就不可能检验如:$H_0:\mu=\mu_0$ 的参数假设,然而还可用非参数方法来分析数据的其他方面,这并不需要严格的假定(见第二十章).

虽然术语"参数假设"用与假设检验有关的,但是第十四、十五章中的用于估计参数估计方法,也是参数统计推断的组成部分.为得到参数的合理的置信区间,每种方法都有特殊的、严格的假定(需要).

16.26 检验正态分布总体方差 σ^2 的假设

关于总体方差的假设可用单样本检验,其中检验的零假设为 $H_0:\sigma^2=\sigma_0^2$,对立假设为下述三个之一(见16.3节):$H_1:\sigma^2>\sigma_0^2$,$H_0:\sigma^2<\sigma_0^2$ 或 $H_0:\sigma^2\neq\sigma_0^2$.我们已知样本方差 S^2 是 σ^2 的最小方差无偏估计(见14.2节),且每一特定样本估计值 $S^2=s^2$ 是 σ^2 的一个点估计值.由

15.2 节还知道总体方差合适的检验统计量(当有一单样本且总体为正态分布)是用 σ_0^2 代替 σ^2 的 χ^2 统计量[(15.1)式]:

$$X^2 = \frac{(n-1)S^2}{\sigma_0^2} \tag{16.13}$$

可假定的特定值:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \tag{16.14}$$

χ^2 统计量满足检验统计量的两个标准(见 16.3 节):(1)能将点估计值(s^2)与零假设值(σ_0^2)作比较;(2)在假定 H_0 为真下,有已知的概率分布($\nu = n-1$ 个自由度的 χ^2 分布).

我们用 χ^{2*} 表示给定样本算得的 X^2 特定值.因此,用 X^2 为检验统计量的假设检验的 P 值是观测的条件概率:即当 H_0 为真时,在自由度($\nu=n-1$)规定的某个 χ^2 分布中,观测到至少与 χ^{2*} 一样的 χ^2 值.对上述三组统计假设对有如下叙述:

对 $H_0:\sigma^2=\sigma_0^2$ 和 $H_1:\sigma^2>\sigma_0^2$ 的右尾检验,P 值为

$$P = P(X^2 \geqslant \chi^{2*} \mid H_0 \text{为真}) \tag{16.15}$$

对 $H_0:\sigma^2=\sigma_0^2$ 和 $H_1:\sigma^2<\sigma_0^2$ 的左尾检验,P 值为

$$P = P(X^2 \leqslant \chi^{2*} \mid H_0 \text{为真}) \tag{16.16}$$

对 $H_0:\sigma^2=\sigma_0^2$ 和 $H_1:\sigma^2\neq\sigma_0^2$ 的双侧检验,因为 X^2 分布不对称,仅取正值,故不能直接给出以概率陈述一个可能的解法是下列步骤:

(1) 确定 χ^{2*} 值是大于还是小于 χ^2 分布的均值 $\mu=\nu=n-1$(见 15.2 节).

(2) 若 $\chi^{2*}>\mu$,则 P 值为

$$P = 2[P(X^2 \geqslant \chi^{2*} \mid H_0 \text{为真})] \tag{16.17}$$

(3) 若 $\chi^{2*}<\mu$,则 P 值为

$$P = 2[P(X^2 \leqslant \chi^{2*} \mid H_0 \text{为真})] \tag{16.18}$$

同 t 统计量一样(见 16.19 节),用 X^2 统计量一般也仅能由表 A.7(**χ^2 分布临界值**)得出近似的或界定的 P 值.因此,我们仅介绍临界值决策规则.同样也和 t 统计量一样,因为 X^2 统计量的临界值依赖于 α 和($\nu=n-1$),所以我们这里只给出决策规则的一般形式.

对于右尾检验:

$$\text{拒绝 } H_0,\quad \text{若 } \chi^{2*} > \chi^2_{\alpha,\nu}$$

对于左尾检验:

$$\text{拒绝 } H_0,\quad \text{若 } \chi^{2*} < \chi^2_{1-\alpha,\nu}$$

对于双侧检验:

$$\text{拒绝 } H_0,\quad \text{若 } \chi^{2*} > \chi^2_{\alpha/2,\nu} \text{或} \chi^{2*} < \chi^2_{\alpha/2,\nu}$$

例 16.9　一个样本($n=16$)取自 σ^2 未知的正态分布总体,若样本方差 $s^2=2.86$.在 $\alpha=0.01$ 下用临界值决策规则给出 $H_0:\sigma^2=2.50$ 的右尾检验,用 16.16 节的步骤给出解.

解　(1)$H_0:\sigma^2=2.50$,$H_1:\sigma^2>2.50$;

(2) $\alpha=0.01$;

(3) 因为总体是正态分布,所以用 X^2 统计量[(16.12)式].对 $\alpha=0.01$,($\nu=n-1=16-1=15$)以及由表 A.7 得到的 $\chi^2_{0.01,15}$ 值,右尾检验决策规则为:拒绝 H_0,若 $\chi^{2*}>(\chi^2_{0.01,15}=30.58)$.

(4) X^2 统计量的值为

$$\chi^{2*}=\frac{(n-1)s^2}{\sigma_0^2}=\frac{(16-1)2.86}{2.50}=17.160,\text{或 }17.16$$

由表 A.7 可见对于 $\nu=15$ 的分布，χ^{2*} 值在 $\chi^2_{0.50,15}$ 和 $\chi^2_{0.10,15}$ 之间. 于是由界定法，P 值范围是 $0.10<P<0.50$.

(5) 因为 $\chi^{2*}<30.58$，所以接受 H_0. 由 $P>0.01$ 可确认这一结果.

16.27 检验正态分布总体标准差 σ 的假设

在研究总体标准差 σ 时，零假设为 $H_0:\sigma=\sigma_0$，三个对立假设为 $H_0:\sigma>\sigma_0$，$H_1:\sigma<\sigma_0$ 和 $H_1:\sigma\neq\sigma_0$. 由于零假设说"σ 等于某个 σ_0"，这意味着 σ 的平方与该值的平方，即方差相等. 因而 16.26 节中相同的检验统计量 X^2 和决策方法也适用于此.

例 16.10 一个样本($n=28$)来自于标准差 σ 未知的正态分布总体. 若样本标准差 $s=4.91$，在 $\alpha=0.05$ 下，用临界值决策规则给出 $H_0:\sigma=3.50$ 的右尾检验，并用 16.16 节中的步骤给出解.

解 (1) $H_0:\sigma=3.50$，$H_1:\sigma>3.50$；

(2) $\alpha=0.05$；

(3) 因为总体是正态分布，所以用 X^2 统计量[(16.13)式]. 由 $\alpha=0.05$，($\nu=n-1=28-1=27$)以及由表 A.7 得到的 $\chi^2_{0.05,27}$ 值，右尾决策规则为：拒绝 H_0，若 $\chi^{2*}>(\chi^2_{0.05,27}=40.11)$.

(4) X^2 统计量的值为

$$\chi^{2*}=\frac{(n-1)s^2}{\sigma_0^2}=\frac{(28-1)(4.91)^2}{(3.50)^2}=53.136,\text{或 }53.14$$

由表 A.7 可见对于 $\nu=27$ 的分布，此 χ^{2*} 值比 $\chi^2_{0.005,27}$ 大，于是 P 值范围是：$P<0.005$.

(5) 因为 $\chi^{2*}>40.11$，所以拒绝 H_0 接受 H_1. $P<0.05$ 可确认这一结果.

16.28 检验二项总体比率 p 的假设：大样本

关于二项总体比率 p 的假设可由近似正态方法或二项分布方法来检验，方法的选择依赖于样本的大小. 本节我们考虑大样本方法，在 16.29 节中讨论小样本方法.

在 Bernoulli 试验条件下，考虑取自无限大二项分布总体的 n 个元素的样本(见 13.21)，其中成功元素的比率 p 未知. 如果假设比率值为 p_0，则应该用什么样的统计量来检验 $H_0:p=p_0$ 与三种对立假设 $H_1:p>p_0$、$H_1:p<p_0$ 或 $H_0:p\neq p_0$ 之一呢?

由 13.25 节知，若 $H_0:p=p_0$ 为真，且 n"足够大"($np_0\geqslant5$ 且 $nq_0\geqslant5$)，则比率的抽样分布是近似正态分布，$\mu_{\bar{p}}=p_0$，$\sigma_{\bar{p}}=\sqrt{\frac{p_0(1-p_0)}{n}}$，因此在这些条件下，用 p_0 代替 $\mu_{\bar{p}}$ 的随机变量 Z[(13.30)式]为

$$Z=\frac{\bar{P}-\mu_{\bar{p}}}{\sigma_{\bar{p}}}=\frac{\bar{P}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}\tag{16.19}$$

假定特定值

$$z=\frac{\bar{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}\tag{6.20}$$

有近似标准正态分布. 于是，对比率假设的大样本检验，这个符合 16.3 节中两个标准的 Z 统计量是合适的检验统计量，因此，对 Z 统计量确定 P 值(见 16.6 节)和决策规则(见 16.10 节)的过程都适用于此.

例 16.11 在 Bernoulli 试验条件下，一个 $n=100$ 的样本取自的无限大二项分布，其中成功的元素比率 p

未知,且得到 $\bar{p}=0.49$.在 $\alpha=0.05$ 下,用临界值决策规则给出 $H_0:p=0.40$ 的右尾检验,并以16.16节中步骤给出解.

解 ☞ (1) $H_0:p==0.40, H_1:p>0.40$;

(2) $\alpha=0.05$;

(3) 因为 $(np_0=100\times0.40=40)>5$,且 $(nq_0=100\times0.60=60)>5$,所以用近似标准正态分布的 Z 统计量[(16.19)式].对 $\alpha=0.05$,右尾检验决策规则为:拒绝 H_0,若 $z^*>1.645$.

(4) Z 统计量值为

$$z=\frac{\bar{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}=\frac{0.49-0.40}{\sqrt{\frac{0.40(1-0.40)}{100}}}=1.837, \text{或 } 1.84$$

由表A.7得近似值:

$$P\approx P[Z\geqslant(z^*=a)\mid H_0\text{ 为真}]=P(Z>1.84)=0.5-0.4671=0.0329$$

(5) 因为 $z^*>1.645$,所以拒绝 H_0 接受 H_1.由近似 P 值小于0.05可确认该结论.

16.29 检验二项总体比率 p 的假设:小样本

当小样本取自二项总体,且 np_0 和 nq_0 都小于5,则不能应用16.28节中的近似正态方法.我们用样本中成功总数的二项随机变量 Y 作为检验统计量,变量由(13.22)式定义为

$$Y=n\bar{P} \tag{16.21}$$

可假定特定值

$$y=n\bar{p} \tag{16.22}$$

若 H_0 为真,Y 有已知的概率分布—二项分布(见15.5节),我们可以直接由此分布确定假设检验中精确 P 值.

例 16.12 若认为一硬币得到正面的比率比反面大,投掷七次硬币得到六次为正面.如果以得到正面为成功,则此硬币的无限二项总体有未知的成功元素比例 p.由此例对 $\alpha=0.05$ 下 $H_0:p=0.5$ 进行右尾检验,并按16.16节中步骤给出解.

解 ☞ (1) $H_0:p=0.5,\ H_1:p>0.5$;

(2) $\alpha=0.05$;

(3) 因为 $(np_0=7\times0.5=3.5)<5$ 且 $(nq_0=7\times0.5=3.5)<5$,所以用 Y 作为检验统计量.右尾决策规则为:拒绝 H_0,若 $P\leqslant0.5$.用记号 y^* 来表示 Y 的特定样本值,则可由此式直接计算 P 值:

$$P=P(Y\geqslant y^*\mid H_0\text{ 为真})$$

(4) 对此样本 $y^*=6$,且由表11.2知

$$P=P(Y\geqslant6\mid H_0\text{ 为真})=0.0547+0.0078=0.0625$$

(5) 因为 $P>0.05$,所以接受 H_0.在0.05的显著水平下,没有证据表明正面的比率比反面大.

习题解答

总体均值 μ 的假设检验:已知标准差 σ 的正态分布总体

16.1 销售胡萝卜种子的公司对硬黏土进行了一年质量控制检验,除了别的以外是否仍可以说此种子的胡萝卜平均长度为11.5cm(见12.8节).现取40个长成的胡萝卜样本,得其平均长度为 $\bar{x}=11.8$cm.假定长度测量值总体仍是 $\sigma=1.15$cm 的正态分布,在显著水平 $\alpha=0.05$ 下,用临界值决策规则(见16.10节)给出 $H_0:\mu=11.5$cm 的双尾检验,

并按 16.16 中所列步骤给出解.

解 ☞ (1) $H_0:\mu=11.5\text{cm}$, $H_1:\mu\neq11.5\text{cm}$;

(2) $\alpha=0.05$;

(3) 因为总体是正态分布的且 σ 已知,于是可用 Z 统计量[(16.2)式].在 $\alpha=0.05$ 下,双边检验决策规则为:拒绝 H_0,若 $z^*>1.96$ 或 $z^*<-1.96$.

(4) Z 统计量值为

$$z^*=\frac{\bar{x}-\mu_0}{\sigma_{\bar{x}}}=\frac{11.8\text{cm}-11.5\text{cm}}{1.15\text{cm}/\sqrt{40}}=1.6499 \text{ 或 } 1.65$$

为确定 P 值,由(16.4)式:

$$P=2[P(Z\geqslant< z^*=a>|H_0 \text{ 为真})]$$

由表 A.5 知,标准正态分布在以上区间 $0\leqslant z\leqslant1.65$ 的面积(概率)为 0.4505,于是

$$P=2(0.5-0.4505)=0.0990$$

(5) 因为 $-1.96<z^*<1.96$,所以接受 H_0,由 $P>0.05$ 可确认这一结果,这些结果提醒公司,由于不能拒绝 $H_0:\mu=11.5\text{cm}$,所以他们应在广告中继续说 $\mu=11.5\text{cm}$.

16.2 有理由相信习题 16.1 中的硬黏土上的胡萝卜不会比沙土中的长得长,即平均值不可能为 11.5cm.将这些胡萝卜种于沙土中,然后取 50 个长成的胡萝卜样本,发现其平均长度为 $\bar{x}=11.1\text{cm}$,假定长度测量值总体仍是 $\sigma=1.15\text{cm}$ 的正态分布.在 $\alpha=0.01$ 下,用临界值决策规则(见 16.10 节)给出 $H_0:\mu=11.5\text{cm}$ 的左尾检验,并将解由 16.16 步骤表示.

解 ☞ (1) $H_0:\mu=11.5\text{cm}$, $H_1:\mu<11.5\text{cm}$;

(2) $\alpha=0.01$;

(3) 用 $\alpha=0.01$ 下的 Z 统计量.左尾决策规则为:拒绝 H_0,若 $z^*<-2.33$.

(4) Z 统计量值为

$$z^*=\frac{\bar{x}-\mu_0}{\sigma_{\bar{x}}}=\frac{11.1\text{cm}-11.5\text{cm}}{1.15\text{cm}/\sqrt{50}}=-2.4595 \text{ 或 } -2.46$$

为确定 P 值,由(16.5)式等价形式:

$$P=P(Z<-2.46\,|\,H_0 \text{ 为真})$$

用例 16.1 中方法及由表 A.5 得的近似值,于是 P 值,

$$P=0.5-0.4931=0.0069$$

(5) 因为 $z^*<-2.33$,所以拒绝 H_0 接受 H_1.这一结果可由 $P<0.01$ 予以确认.因此在 0.01 显著水平下,有证据认为沙土中胡萝卜长度小于 11.5cm.

16.3 由例 16.1 和 16.2 以及习题 16.1 和 16.2 给出的统计决策,这些决策下可能产生什么样的错误,第 I 类还是第 II 类?

解 ☞ 在例 16.1 和 16.2 以及习题 16.2 中,拒绝 H_0,于是可能产生第一类错误(见表 16.1).在习题 16.1 中,接受 H_0,于是可能产生第 II 类错误.

16.4 重复习题 16.1 的双侧检验,但这次用 $\bar{X}$ 作为检验统计量,按 16.16 节的步骤给出解.

解 ☞ (1) $H_0:\mu=11.5\text{cm}$, $H_1:\mu\neq11.5\text{cm}$;

(2) $\alpha=0.05$;

(3) 因为总体是正态分布且 σ 已知,所以我们可以用 $\bar{X}$ 作为检验统计量(见 16.17 节).对 $\alpha=0.05$,由表 A.5 得:$z_{\alpha/2}=z_{0.05/2}=1.96$,以临界值叙述双边决策规则如下:拒绝 H_0,如果

$$\bar{x}^*>\left[\bar{x}_{0.05/2}=\mu_0+z_{0.05/2}\sigma_{\bar{x}}=11.5\text{cm}+1.96\left(\frac{1.15\text{cm}}{\sqrt{40}}\right)=11.86\text{cm}\right]$$

或

$$\bar{x}^* < \left[\bar{x}_{1-0.05/2} = \mu_0 - z_{0.05/2}\sigma_{\bar{x}} = 11.5\text{cm} - 1.96\left(\frac{1.15\text{cm}}{\sqrt{40}}\right) = 11.14\text{cm}\right]$$

(4) $\bar{X}$ 统计量的值为 $\bar{x}^* = 11.8$cm，P 值与习题 16.1 中的 $P = 0.0990$ 相同.

(5) 由于 $11.14\text{cm} < \bar{x}^* < 11.86\text{cm}$，所以仍接受 H_0. 这一结果可由 $P > 0.05$ 予以确认.

β，$1-\beta$，功效曲线及 OC 曲线

16.5　在例 16.5 中，对于 $n = 20$，$\sigma = 6.25$min 及 $\alpha = 0.05$，我们发现如果 $\mu_1 = 105.3$min 是真正的总体均值，则右尾检验拒绝 H_0 的概率为 $1 - \beta(105.3) = 0.9732$. 在相同条件下，如果真正的总体均值为：(a)101.3min，(b)100.3min，(c)99.3min 时，拒绝 H_0 的概率为多少？

解　(a)图 16-7 给出了问题的条件. 其中下面的是零假设定义的分布与图 16-5 中的一样，而由 $H_1: \mu = \mu_1 = 101.3$min 定义的上边分布有左移，使得其大部分(灰色区域)位于检验的接受域. 为得出 $\beta(101.3)$，首先来确定 z_1 值：

$$z_1 = \frac{102.6\text{min} - 101.3\text{min}}{6.25\text{min}/\sqrt{20}} = 0.930 \text{ 或 } 0.93$$

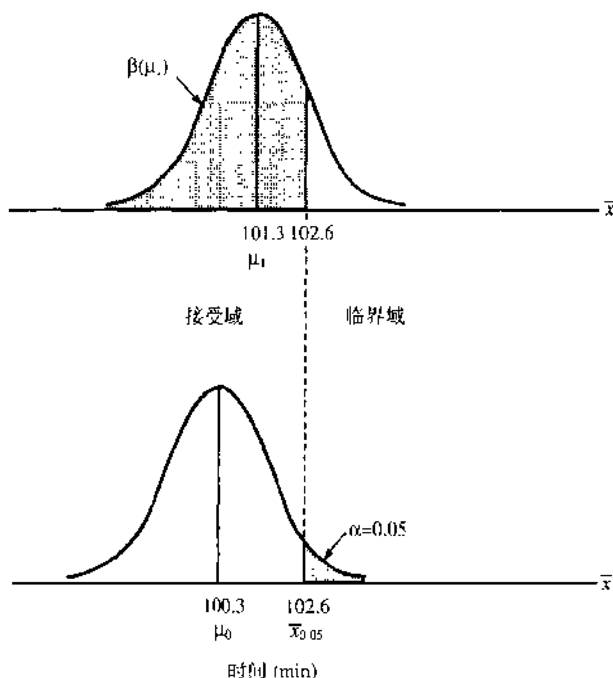


图 16-7

因此由表 A.5，第 II 类错误的概率为

$$\beta(101.3) = 0.5 + 0.3238 = 0.8238$$

于是拒绝 H_0(正确地)概率为

$$1 - \beta(101.3) = 1 - 0.8238 = 0.1762$$

(b) 此时真正均值即是由零假设定义的检验分布的均值. 于是拒绝 H_0(不正确地，第 I 类错误)的概率是图 16-7 中下面分布的黑色区域，检验的显著水平 $\alpha = 0.05$.

(c)此时真实的均值 $\mu_1 = 99.3$min 小于零假设定义的均值 $\mu = 100.3$min，于是拒绝 H_0 而接受 $H_1: \mu > 100.3$min 也是第 I 类错误. 然而现在可从图 16-8 中的上面分布发现，在检验的临界域中真实分布的百分数小于 α，于是第 I 类错误的概率小于 α. 这一问题可说明为：在**单侧检验中，实际上 α 是犯第I类错误的最大可能. 这一最大值仅当真实均值为零假设定义的 μ_0 时达到**. 如果是右尾检验

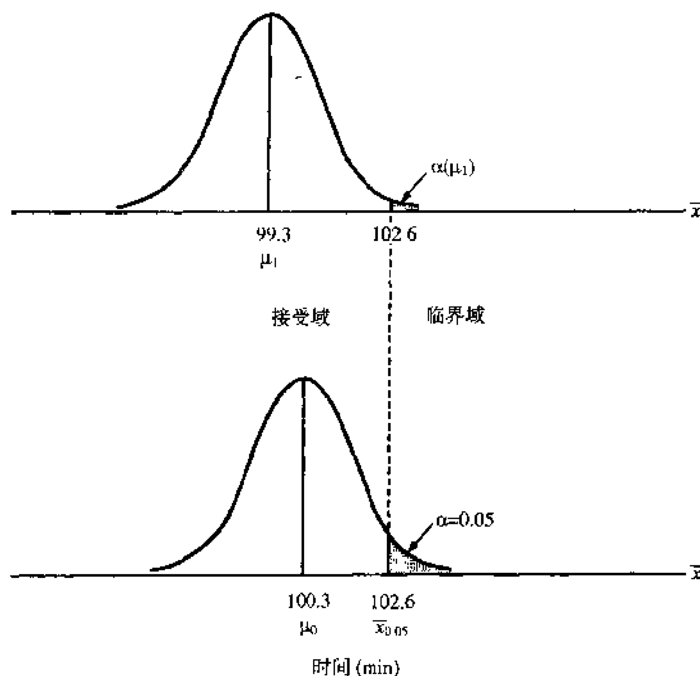


图 16-8

$\mu<\mu_0$,则第Ⅰ类错误的概率小于 α,记为 $\alpha(\mu_1)$.此时为得出 $\alpha(\mu_1)=\alpha(99.3)$的值,我们首先确定 z_1 值:

$$z_1=\frac{102.6\text{min}-99.3\text{min}}{6.25\text{min}/\sqrt{20}}=2.361 \text{ 或 } 2.36$$

然后由表 A.5 得

$$\alpha(99.3)=0.5-0.4909=0.0091$$

因此,拒绝 H_0(不正确地,犯第Ⅰ类错误)的概率为 0.0091.

16.6 在例 16.4 和 16.5 及习题 16.5 中,我们对不同的总体均值 μ_λ 得到了接受或拒绝 $H_0:\mu=100.3$min的概率,这些概率已列在表 16.3.在同样条件下,仍假定同样的右尾检验,完成此表.

表 16.3

实际均值(min)	97.3	98.3	99.3	100.3	101.3
P(接受 H_0)	$1-\alpha(97.3)$ =	$1-\alpha(98.3)$ =	$1-\alpha(99.3)$ =0.9909	$1-\alpha$ =0.95	$\beta(101.3)$ =0.8238
P(拒绝 H_0)	$\alpha(97.3)$ =	$\alpha(98.3)$ =	$\alpha(99.3)$ =0.0091	α =0.05	$1-\beta(101.3)$ =0.1762
实际均值(min)	102.3	103.3	104.3	105.3	106.3
P(接受 H_0)	$\beta(102.3)$ =	$\beta(103.3)$ =	$\beta(104.3)$ =	$\beta(105.3)$ =0.0268	$\beta(106.3)$ =
P(拒绝 H_0)	$1-\beta(102.3)$ =	$1-\beta(103.3)$ =	$1-\beta(104.3)$ =	$1-\beta(105.3)$ =0.9732	$1-\beta(106.3)$ =

解 ☞　完成的表由表 16.4 给出.

表 16.4

实际均值(min)	97.3	98.3	99.3	100.3	101.3
P(接受 H_0)	$1-\alpha(97.3)$ $=0.9999$	$1-\alpha(98.3)$ $=0.9990$	$1-\alpha(99.3)$ $=0.9909$	$1-\alpha$ $=0.95$	$\beta(101.3)$ $=0.8238$
P(拒绝 H_0)	$\alpha(97.3)$ $=0.0001$	$\alpha(98.3)$ $=0.0010$	$\alpha(99.3)$ $=0.0091$	α $=0.05$	$1-\beta(101.3)$ $=0.1762$
实际均值(min)	102.3	103.3	104.3	105.3	106.3
P(接受 H_0)	$\beta(102.3)$ $=0.5832$	$\beta(103.3)$ $=0.3085$	$\beta(104.3)$ $=0.1112$	$\beta(105.3)$ $=0.0268$	$\beta(106.3)$ $=0.0040$
P(拒绝 H_0)	$1-\beta(102.3)$ $=0.4168$	$1-\beta(103.3)$ $=0.6915$	$1-\beta(104.3)$ $=0.8888$	$1-\beta(105.3)$ $=0.9732$	$1-\beta(106.3)$ $=0.9960$

16.7 对所有条件都相同的右尾检验,但样本容量 n 从 20 加倍到 40,再完成表 16.4.

解答 完成的表由表 16.5 给出. 为说明如何得到这些值,我们将对真值 99.3min 和 101.3min 予以计算. 计算的第一步是用例 16.3 中的过程,对均值的零假设定义的抽样分布,确定一个新的临界值 $\bar{x}_{0.05}$,于是

$$\bar{x}_{0.05}=\mu_0+1.645\sigma_{\bar{x}}=100.3\text{min}+1.645\left(\frac{6.25\text{min}}{\sqrt{40}}\right)=101.93\text{min}$$

下面用习题 16.5(c)中的方法求 $\alpha(99.3)$,且首先得出

$$z_1=\frac{101.93\text{min}-99.3\text{min}}{6.25\text{min}/\sqrt{40}}=2.661,\text{或}\ 2.66$$

然后,由表 A.5 得

$$\alpha(99.3)=0.5-0.4961=0.0039$$

于是

$$1-\alpha(99.3)=1-0.0039=0.9961$$

表 16.5

实际均值(min)	97.3	98.3	99.3	100.3	101.3
P(接受 H_0)	$1-\alpha(97.3)$ $=1$	$1-\alpha(98.3)$ $=0.9999$	$1-\alpha(99.3)$ $=0.9961$	$1-\alpha$ $=0.95$	$\beta(101.3)$ $=0.7389$
P(拒绝 H_0)	$\alpha(97.3)$ $=0.0000$	$\alpha(98.3)$ $=0.0001$	$\alpha(99.3)$ $=0.0039$	α $=0.05$	$1-\beta(101.3)$ $=0.2611$
实际均值(min)	102.3	103.3	104.3	105.3	106.3
P(接受 H_0)	$\beta(102.3)$ $=0.3557$	$\beta(103.3)$ $=0.0823$	$\beta(104.3)$ $=0.0082$	$\beta(105.3)$ $=0.0003$	$\beta(106.3)$ $=0.0000$
P(拒绝 H_0)	$1-\beta(102.3)$ $=0.6443$	$1-\beta(103.3)$ $=0.9177$	$1-\beta(104.3)$ $=0.9918$	$1-\beta(105.3)$ $=0.9997$	$1-\beta(106.3)$ ≈ 1

为得出 $\beta(101.3)$值,我们用习题 16.5(a)中方法,首先确定

$$z_1=\frac{101.93\text{min}-101.3\text{min}}{6.25\text{min}/\sqrt{40}}=0.638\ \text{或}\ 0.64$$

查表 A.5 得

$$\beta(101.3)=0.5+0.2389=0.7398$$

于是

$$1-\beta(101.3)=0.5-0.7389=0.2611$$

16.8 由表16.4和16.5,在同一坐标系下给出 $n=20$ 和 $n=40$ 的功效曲线(见16.18节).

解 ☞ 所求功效曲线由图16-9给出.横轴是真实的总体均值 μ_r 的可能取值,纵轴是拒绝零假设 $H_0:\mu=100.3\text{min}$ 的概率.为了完整性,这些曲线给出了当 H_0 应被拒绝($\mu_r>100.3\text{min}$)和接受($\mu_r\leqslant 100.3\text{min}$)的拒绝概率.当拒绝 H_0 时,从横轴上特定 μ_r 点到检验曲线上的垂直距离为该检验功效(见16.18节)$1-\beta(\mu_r)$.于是当拒绝 H_0 时,对任一给定 μ_r,第II类错误的概率 $\beta(\mu_r)$ 是在 μ_r 上面从功效曲线到1.0线的距离,$\beta(101.3)$ 如图16-9所示.当 H_0 应被接受时,由横轴上特定 μ_r 点到检验曲线的垂直距离是该检验犯第I类错误概率 $\alpha(\mu_r)$,其中 $\alpha=0.05$,$\mu_r=\mu_0=100.3\text{min}$,于是在应接受 H_0 的地方,即对某一特定 μ_r,从检验的功效曲线到1.0线的垂直距离,$1-\alpha(\mu_r)$,是检验正确接受 H_0 的概率.

图16-9给出的两条曲线是针对下列检验条件的:右尾检验,$\alpha=0.05$,$n=20$ 或40,$\bar{X}$ 的正态抽样分布.这样的**S型曲线**是典型的右尾检验功效曲线,而**反S型曲线**是典型的左尾检验功效曲线.双侧检验的功效曲线是两种曲线类型的组合.

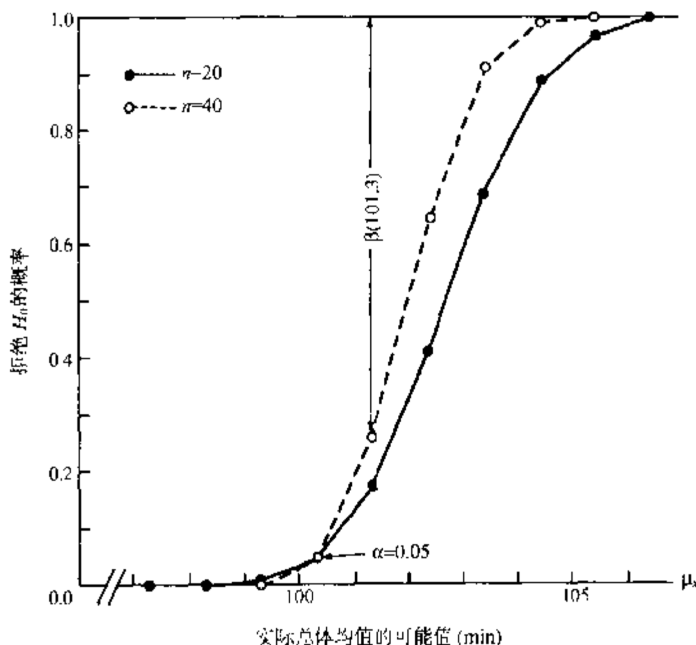


图16-9

16.9 由表16.4和16.5就 $n=20$ 和 $n=40$ 情况,在同一坐标下绘出OC线(操作特性曲线,见16.18节).

解 ☞ 图16-10给出所求的OC曲线.和图16-9的功效曲线一样,横轴表示真正总体均值 μ_r 的可能值,而现在的纵轴表示接受零假设 $H_0:\mu=100.3\text{min}$ 的概率.同样,为了完整性,这些曲线给出了拒绝和接受 H_0 时的接受概率.这些曲线是图16-9中的曲线的补,因此当拒绝 H_0 时,由横轴上任一点 μ_r 到曲线的垂直距离为 $\beta(\mu_r)$,而从交点到1.0线的垂直距离为 $1-\beta(\mu_r)$.$1-\beta(103.3)$ 如图16.10所示.类似地,当接受 H_0 时,由横轴上任一点 μ_r 到曲线的垂直距离为 $1-\alpha(\mu_r)$

同样也如图16-9,曲线是针对这些右尾条件的.这样条件下的OC曲线是典型的反S型,对左尾检验OC曲线应为S型.同样,双侧检验的OC曲线是这些曲线的组合.

16.10 样本容量的变化对第I、第II类错误的概率有何影响?

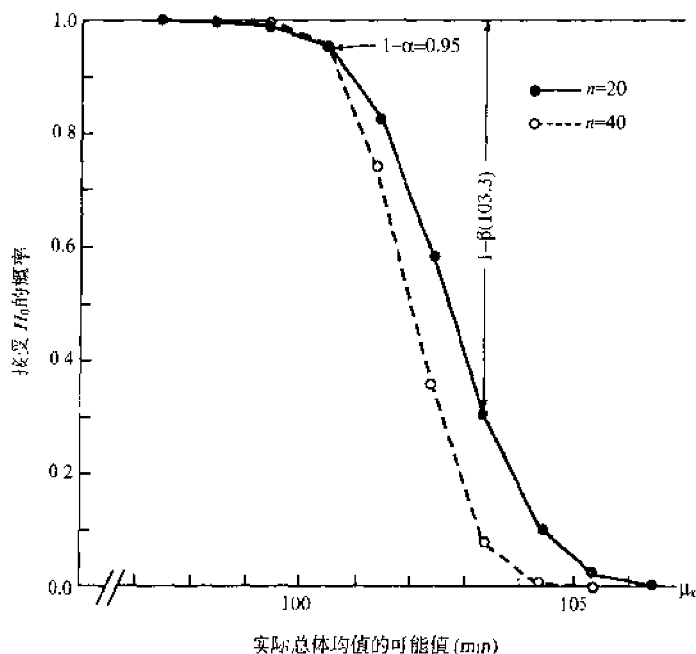


图 16-10

解 由习题 16.6 到 16.9 中的图表和曲线可见,若其他检验条件保持不变,仅增加样本容量,两种错误类型的概率 $\alpha(\mu_x)$ 和 $\beta(\mu_x)$ 同时减少,当 α 保持常数时,这对所有不是 μ_0 的 μ_x 都成立.

16.11 在备选的 μ_x 和零假设定义的 μ_0 间距离的变化对第 I、第 II 类错误概率有何影响?

解 由习题 16.6 到 16.9 中的图表和曲线可知,若其他检验条件保持不变,而 μ_x 与 μ_0 间距离增加,则在右尾检验中,如果 μ_x 在 μ_0 左侧,$\alpha(\mu_x)$ 将减小,如果 μ_x 在 μ_0 右侧,$\beta(\mu_x)$ 将减小.

16.12 减小 α,对第 I、第 II 类错误的概率如何变化?

解 若其他检验条件保持不变,仅 α 减小,则对所有应接受的 H_0 的 μ_x,$\alpha(\mu_x)$ 将减小.而对所有应拒绝 H_0 的 μ_x,$\beta(\mu_x)$ 将增大.为避免随 α 减小,$\beta(\mu_x)$ 增大的唯一途径就是增加样本大小.

总体均值的假设检验:来自标准差未知正态分布总体的小样本

16.13 对于 $n=25$ 的样本,$H_0:\mu=\mu_0$ 的单样本假设检验的决策规则是什么?其中以 t 为检验统计量,$\alpha=0.05$ 及(a) $H_1:\mu>\mu_0$,(b) $H_1:\mu<\mu_0$ (c) $H_1:\mu\neq\mu_0$.

解 因为用表 A.6 仅能求出近似或界定 P 值(见 16.20 节),所以我们用临界值决策规则(见 16.21 节):

(a) 对右尾检验,其中 $\nu=n-1=25-1=24$:

$$\text{拒绝 } H_0,\quad \text{若 } t^*>(t_{\alpha,\nu}=t_{0.05,24}=1.711)$$

(b) 对左尾检验,其中 $\nu=24$,

$$\text{拒绝 } H_0,\quad \text{若 } t^*<(-t_{\alpha,\nu}=-t_{0.05,24}=-1.711)$$

(c) 对双边检验,其中 $\nu=24$,

$$\text{拒绝 } H_0,\text{若 } t^*>(t_{\alpha/2,\nu}=t_{0.05/2,24}=2.064)\text{ 或 } t^*<(-t_{\alpha/2,\nu}=-t_{0.05/2,24}=-2.064)$$

16.14 对习题 16.13 的检验条件,若 $\alpha=0.01$ 决策规则又是什么?

解 (a) 对右尾检验,其中 $\nu=24$,有

$$\text{拒绝 } H_0,\quad \text{若 } t^*>(t_{\alpha,\nu}=t_{0.01,24}=2.492)$$

(b) 对左尾检验,其中 $\nu=24$,有

$$\text{拒绝 } H_0,\quad \text{若 } t^*<(-t_{\alpha,\nu}=-t_{0.01,24}=-2.492)$$

(c) 对双边检验,其中 $\nu=24$,有

$$拒绝 H_0,若 t^* > (t_{\alpha/2,\nu} = t_{0.01/2,24} = 2.797) 或 t^* < (-t_{\alpha/2,\nu} = -t_{0.01/2,24} = -2.797)$$

16.15 一位研究大平原硬甲虫的生物学家发现生活在3000英尺高山上的该种类的总体.从中取 $n=20$ 个雄甲虫的不同的度量种类,以考察高山上的雄甲虫是否不同于平原上的雄甲虫,其中度量之一是翅膀上黑斑的长度.已知平原雄甲虫黑斑长度是 $\mu=3.14$mm的正态分布,从高山雄甲虫样本得到黑斑长度为 $\bar{x}=3.23$mm,$s=0.4$mm.假定高山雄甲虫斑长为正态分布,对 $\alpha=0.05$ 做 $H_0:\mu=3.14$mm的双边检验,用16.16节的步骤给出解.

解 ☞ (1) $H_0:\mu=3.14$mm, $H_1:\mu\neq3.14$mm;

(2) $\alpha=0.05$;

(3) 用 t 统计量[(16.8)式],对 $\alpha=0.05$,($\nu=n-1=20-1=19$),双边检验决策规则(见16.21节)为

$$拒绝 H_0,若 t^* > (t_{0.05/2,19} = 2.093) 或 t^* < (-t_{0.05/2,19} = -2.093)$$

(4) t 统计量的值为

$$t^* = \frac{3.23\text{mm} - 3.14\text{mm}}{0.214\text{mm}/\sqrt{20}} = 1.8808 或 1.881$$

对于 $\nu=19$ 的 P 的值是在区间(见16.20节)

$$(2\times0.025) < P < (2\times0.05) 或 0.05 < P < 0.10$$

(5) 因为 $-2.093<t^*<2.093$,所以接受 H_0.由 P 值落入范围 $0.05<P<0.10$ 中确认这一结果,于是在0.05的显著水平下,样本的平均斑点长度和假设平均斑点长度之间没有显著不同.

16.16 在做习题16.15检验前,假设高山雄甲虫由于吸收热量较多,因此斑点较长,因此你需要进行 $H_0:\mu=3.14$mm 的右尾检验,在 $\alpha=0.05$ 下做此检验,并用16.16节中的步骤给出解.

解 ☞ (1) $H_0:\mu=3.14$mm, $H_1:\mu>3.14$mm;

(2) $\alpha=0.05$

(3) 用 t 统计量,当 $\nu=19$,$\alpha=0.05$ 时,右尾决策规则为:拒绝 H_0,若 $t^*>1.792$.

(4) 同样有 $t^*=1.881$.但现在对于右尾检验 $\nu=19$ 的 P 值在范围 $0.025<P<0.05$ 中.

(5) 因为 $t^*>1.792$,所以拒绝 H_0 接受 H_1,这一结论可由 P 值在范围 $0.025<P<0.05$ 中来确认.在此可以说在0.05显著水平下,样本的平均斑点长比假设的平均斑点长显著的长,而且这些结果表明高山上的甲虫有更长的平均斑点长度.

16.17 标准的牙科麻醉剂在注射后平均 $\mu=10.5$min 后会消除感觉.现有一新型麻醉药品,制造商声称其作用"显著地快".如果你是一个牙医,并对十名患者施用了此药,并得到从注射到失去感觉的下列时间(以分钟计)为:9.3,9.5,9.2,9.0,9.3,9.5,9.4,9.3,9.2,9.1.由于以前生效时间是正态分布,所以合理地假定新麻醉药为正态分布,因为关心的仅是药效时间的减少,用样本信息和 t 统计量在 $\alpha=0.01$ 下对 $H_0:\mu=10.5$min进行左尾检验,并用16.16节中的步骤给出解.

解 ☞ (1) $H_0:\mu=10.5$min, $H_1:\mu<10.5$min;

(2) $\alpha=0.01$;

(3) 用 t 统计量,对 $\alpha=0.01$ 及($\nu=n-1=10-1=9$),左尾决策规则为

$$拒绝 H_0, \quad 若 t^* < -2.821$$

(4) 对此样本

$$\bar{x} = \frac{\sum x_i}{n} = \frac{92.8\text{min}}{10} = 9.28\text{min}$$

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} = 0.1619\text{min}$$

因此，t 统计量的值为

$$t^* = \frac{9.28\text{min} - 10.5\text{min}}{0.1619\text{min}/\sqrt{10}} = -23.8294 \text{ 或 } -23.829$$

$(\nu=9)$时的 P 值范围为：$P<0.0005$.

(5) 因为 $t^* < -2.821$，所以拒绝 H_0 接受 H_1. 由 P 值落入$P<0.0005$ 的范围可确认此结论. 于是可说在 0.01 显著水平下，有证据表明新麻醉药确实作用更快.

16.18 如果习题 16.17 中样本是(以分钟计)：10.5，10.7，11.0，10.7，10.6，10.7，10.8，10.9，10.5，11.0. 则相同地 $H_0:\mu=10.5\text{min}$ 的左尾检验的结果是什么？

解 ☞ 一个左尾检验在样本均值小于假设的总体均值(见 16.6 节)才有一个显著结果，即拒绝 H_0. 现在样本均值为 $\bar{x}=10.74\text{min}$. 由此没有充分理由拒绝 H_0，新药作用更快，所以必须接受 H_0.

检验正态分布总体方差和标准差的假设

16.19 对于 $n=19$ 的样本，给出 $H_0:\sigma^2=\sigma_0^2$ 的单样本假设检验的临界值决策规则(见 16.26 节)，其中检验统计量为 X^2，$\alpha=0.05$，且(a) $H_1:\sigma^2>\sigma_0^2$，(b) $H_1:\sigma^2<\sigma_0^2$，(c) $H_1:\sigma^2\neq\sigma_0^2$.

解 ☞ (a)对于右尾检验，$\nu=19-1=18$ 时有

$$\text{拒绝 } H_0, \quad \text{若 } \chi^{2*} > (\chi^2_{\alpha,\nu} = \chi^2_{0.05,18} = 28.87)$$

(b) 对于左尾检验，$\nu=18$ 时有

$$\text{拒绝 } H_0, \quad \text{若 } \chi^{2*} < (\chi^2_{1-\alpha,\nu} = \chi^2_{0.95,18} = 9.39)$$

(c) 对于双边检验，$\nu=18$ 时有

$$\text{拒绝 } H_0, \text{若 } \chi^{2*} > (\chi^2_{\alpha/2,\nu} = \chi^2_{0.025,18} = 31.53) \text{ 或 } \chi^{2*} < (\chi^2_{1-\alpha/2,\nu} = \chi^2_{0.975,18} = 8.23)$$

16.20 对于习题 16.15 描述的硬甲虫研究，假定已知平原雄甲虫斑点长是 $\mu=3.14\text{mm}$和 $\sigma^2=0.0505\text{mm}^2$的正态分布. 对于 $n=20$ 的高山雄甲虫样本，得出斑点长的 $\bar{x}=3.23\text{mm}$，$s=0.214\text{mm}$，$s^2=0.0458\text{mm}^2$. 假定高山雄甲虫的斑点长是正态分布，在$\alpha=0.05$下，用临界值决策规则做 $H_0:\sigma^2=0.0505\text{mm}^2$ 的双边检验，并由 16.16 节中的步骤给出解.

解 ☞ (1) $H_0:\sigma^2=0.0505\text{mm}^2$，$H_1:\sigma^2\neq0.0505\text{mm}^2$；

(2) $\alpha=0.05$；

(3) 用 X^2 统计量，在 $\alpha=0.05$ 和$(\nu=n-1=20-1=19)$下，双边检验决策规则(见 16.26 节)为

$$\text{拒绝 } H_0, \text{若 } \chi^{2*} > (\chi^2_{0.05/2,19} = 32.85) \text{ 或 } \chi^{2*} < (\chi^2_{1-0.05/2,19} = 8.91)$$

(4) X^2 统计量的值为[(16.14)式]

$$\chi^{2*} = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(20-1)0.0458\text{mm}^2}{0.0505\text{mm}^2} = 17.2317, \text{或 } 17.23$$

因为 $\chi^{2*}<(\mu=\nu=19)$，所以由表 A.7 得到 P 值是在$(2\times0.1)<P<(2\times0.5)$或 $0.2<P<1.0$ 的范围中(见 16.26 节).

(5) 因为 $8.91<\chi^{2*}<32.85$，所以接受 H_0. 这一结论可由 P 值落在范围 $0.2<P<1.0$ 中来确认. 因此在 $\alpha=0.05$ 显著水平下，没有证据表明两总体有不同的方差.

16.21 对某一车型开发了一种新刹车灯，并认为此灯对于汽车刹车灯的反应时间的均值和方差均减少. 由对现在灯的标准检验知，反应时间(每人一次)的总体是 $\mu=0.33\text{sec}$ 和 $\sigma^2=0.00284\text{sec}^2$ 的正态分布. 由相同的标准检验，取新灯的反应时间的样本($n=51$，每人一次)，并得 $\bar{x}=0.25\text{sec}$ 和 $s^2=0.00182\text{sec}^2$. 假定新灯的时间也是正态分布，在

$\alpha=0.05$下做 $H_0:\sigma^2=0.00284\text{sec}^2$ 的左尾检验,并由 16.16 节的步骤给出解.

解 ☞ (1) $H_0:\sigma^2=0.00284\text{sec}^2$, $H_1:\sigma^2<0.00284\text{sec}^2$;

(2) $\alpha=0.05$;

(3) 由 $\nu=50$ 的 X^2 统计量和表 A.7 得左尾决策规则为(见 16.26 节)

$$\text{拒绝 } H_0, \text{若 } \chi^{2*} < (\chi^2_{0.95,50}=34.76)$$

(4) X^2 统计量的值是[(16.14)式]

$$\chi^{2*}=\frac{(51-1)0.00182\text{sec}^2}{0.00284\text{sec}^2}=32.0423 \text{ 或 } 32.04$$

由表 A.7, $\nu=50$ 时的 P 值在范围 $0.01<P<0.025$ 内.

(5) 因为 $\chi^{2*}<34.76$,所以拒绝 H_0 接受 H_1,由 P 值落入范围 $0.01<P<0.025$ 中的事实可对此结论予以确认.因此可以说在 0.05 显著水平下,有明显的证据表明新灯反应时间有较小方差.

16.22 在 355 毫升(ml)罐装苏打饮料上印制的"营养组成"中标明了仅含有 35 毫克的钠.为表明苏打中钠的含量保持在 $\mu=34.5\text{mg}$ 和 $\sigma=0.24\text{mg}$,在正常的质量控制检验中,从生产线上随机选取了 10 听,且在其他检验中,若样本的标准差明显大于 0.24mg(在 $\alpha=0.05$水平下),则停止生产线,且调整苏打配制程序.如果在一次这样的检验中,得到 $s=0.29\text{mg}$,确定是否有必要调整配料,并按 16.16 节中的步骤给出解.

解 ☞ (1) $H_0:\sigma=0.24\text{mg}$, $H_1:\sigma>0.24\text{mg}$;

(2) $\alpha=0.05$

(3)"σ 等于某一 σ_0 值"的零假设表明 σ 的平方和该值的平方相等.此时的 $H_0:\sigma^2=(0.24\text{mg})^2=0.0576\text{mg}^2$,因此仍用关于方差的 X^2 检验统计量和决策过程(见 16.26 节).由 $\nu=10-1=9$ 及表 A.7,得右尾决策规则为:拒绝 H_0,若 $\chi^{2*}>(\chi^2_{0.05,9}=16.92)$.

(4) X^2 统计量的值为

$$\chi^{2*}=\frac{(n-1)s^2}{\sigma^2-0}=\frac{(10-1)(0.29\text{mg})^2}{(0.24\text{mg})^2}=13.1406 \text{ 或 } 13.14$$

由表 A.7, $\nu=9$ 时的 P 值在 $0.1<P<0.5$ 内.

(5) 因为 $\chi^{2*}<16.92$,所以接受 H_0,由 P 值在 $0.1<P<0.5$ 内确认此结论.在 0.05 显著水平下,没有明显的证据认为样本标准差大于要求的 $\sigma=0.24\text{mg}$,因此不需要调整配料程序.

16.23 对习题 16.22 所给信息,在 $\alpha=0.05$ 下,进行 $H_0:\sigma=0.24\text{mg}$ 的双边检验,并以 16.16 节中的步骤给出解.

解 ☞ (1) $H_0:\sigma=0.24\text{mg}$, $H_1:\sigma\neq0.24\text{mg}$;

(2) $\alpha=0.05$;

(3) 用 X^2 统计量,当 $\alpha=0.05$, $\nu=9$ 时,双边决策规则为(见 16.26 节):拒绝 H_0,若 $\chi^{2*}>19.02$或 $\chi^{2*}<2.70$.

(4) 由习题 16.22 中 $\chi^{2*}=13.14$,且因 $\chi^{2*}>(\mu=n-1=9)$(见 16.26 节),P 值在$(2\times0.1)<P<(2\times0.5)$或 $0.2<P<1.0$ 范围内.

(5) 因为$2.70<\chi^{2*}<19.02$,所以接受 H_0,同样可由 P 值落在范围 $0.2<P<1.0$ 确认这一结论.因此在 0.05 显著水平下,没有明显证据认为需要调整配料程序.

二项分布比率的假设检验

16.24 大学心理学系主任选取了本系 1727 名本科生中的 75 名为随机样本,并询问每名同学本系设立的哪些选修课程应该保留?他决定如果想保留此课程的本科生明显少于 20%,则取消该门课程.结果对课程**心理学**,75 人中有 11 人想保留.在 $\alpha=0.01$ 下,用临界值决策规则对 $H_0:p=0.20$ 进行左尾检验,以确定其命运,并用 16.16 节中步骤给出解.

解 ☞ (1) $H_0:p=0.20$, $H_1:p<0.20$;

(2) $\alpha=0.01$;

(3) 考虑目前本科生是 $N=1727$ 的总体，$n=75/1,727N$ 或 $0.043N$. 于是有 $n<0.05N$，我们可以认为尽管抽样是无放回的，但 p 基本上是不变的(见 11.2 节)，而且不需要用有限总体修正因子乘以 $\sigma_{\bar{p}}$(见 13.11 节). 于是，因为 $(np_0=75\times0.2=15)>5$ 且 $(nq_0=75\times0.8=60)>5$ 所以检验统计量 Z[(16.9)式]具有近似标准正态分布. 由表 A.5 知 $z_{0.01}=2.33$，故决策规则为：拒绝 H_0，若 $z^*<-2.33$.

(4) 因为 $p=0.2$ 且 $\bar{p}=11/75=0.147$，所以 Z 统计量值为

$$z^*=\frac{0.147-0.20}{\sqrt{\frac{0.20(1-0.20)}{75}}}=-1.148,\ \text{或}\ -1.15$$

由表 A.5 得近似值为

$$P\approx P[Z\leqslant(z^*=-a)\quad H_0\ \text{为真}]=P(Z<-1.15)=0.5-0.3749=0.1251$$

(5) 因为 $z^*>-2.33$，所以在 0.01 显著水平下接受 H_0. 用近似 P 值大于 0.01 可以确认这一结论. 没有明显的证据表明对心理学感兴趣的人少于 20%，因此应保留该课程.

16.25 你认为一枚硬币也许是不均匀的，如果不均匀，则测量结果是有较多的正面或有较多的反面，投币七次并发现六次是正面，试在 $\alpha=0.05$ 下，对 $H_0:p=0.5$ 的双尾检验，并用临界值决策规则和 16.16 节步骤给出解.

解 (1) $H_0:p=0.5$, $H_1:p\neq0.5$;

(2) $\alpha=0.01$;

(3) 用 16.29 节中的小样本二项方法，并将 α 平均分到分布两个尾部，则决策规则为(见例 16.12)：拒绝 H_0，若 y^* 是 0 或 7.

(4) $y^*=6$;

(5) 因为 $0<y^*<7$，所以在 5% 显著水平下接近 H_0.

16.26 一制药公司宣称一种新型生发剂可使试用人中的 60% 长出头发. 由得到的反馈，消费者组织认为此比率过高. 该组织将此药用于 8 个相同年龄的人，在一定时间，发现 8 人中只有 3 人长出头发. 就此样本，给出 $\alpha=0.05$ 下 $H_0:P=0.6$ 的左尾检验，并用 P 值决策规则和 16.16 中步骤给出解.

解 (1) $H_0:p=0.6$, $H_1:p<0.6$;

(2) $\alpha=0.05$;

(3) 用 16.29 节中的小样本二项方法，P 值为

$$P=P(Y\leqslant y^*\mid H_0\ \text{为真})$$

且决策规则为：拒绝 H_0，若 $P<0.05$.

(4) 对此样本 $y^*=3$，$H_0:p=0.6$ 的二项分布已由表 16.6 给出. 由此表知：

$$P=P(Y\leqslant3\mid H_0\ \text{为真})=0.123863+0.041288+0.007864+0.000655=0.173670$$

(5) 因为 $P>0.05$，所以接受 H_0. 在 5% 显著水平下没有明显的证据表明所说比率过高.

16.27 对于习题 16.26 中的样本，做 $\alpha=0.05$ 下 $H_0:P=0.6$ 的双边检验，并由临界值决策规则及 16.16 节中的步骤给出解.

解 (1) $H_0:p=0.6$, $H_1:p\neq0.6$;

(2) $\alpha=0.05$;

(3) 如习题 16.26 中，应用二项检验，并将此 0.05 平均分配到分布两个尾部. 由表 16.6，决策规则为：拒绝 H_0，若 y^* 为 0,1 或 8.

(4) $y^*=3$. 对此侧检验，由于分布的对称性，我们不计算 P 值.

(5) 因为 $1<y^*<8$，所以还是接受 H_0.

表 16.6

长头发人数 y	概率 $f(y)=\binom{n}{y}p^yq^{n-y}$
0	$f(0)=\binom{8}{0}(0.6)^0(0.4)^8=0.000655$
1	$f(1)=\binom{8}{1}(0.6)^1(0.4)^7=0.007864$
2	$f(2)=\binom{8}{2}(0.6)^2(0.4)^6=0.041288$
3	$f(3)=\binom{8}{3}(0.6)^3(0.4)^5=0.123863$
4	$f(4)=\binom{8}{4}(0.6)^4(0.4)^4=0.232243$
5	$f(5)=\binom{8}{5}(0.6)^5(0.4)^3=0.278692$
6	$f(6)=\binom{8}{6}(0.6)^6(0.4)^2=0.209019$
7	$f(7)=\binom{8}{7}(0.6)^7(0.4)^1=0.089580$
8	$f(8)=\binom{8}{8}(0.6)^8(0.4)^0=0.016796$

补充习题

总体均值的假设检验

16.28 已知八月中旬雄地鼠的体重总体为：$\mu=400\text{g}$，$\sigma=100\text{g}$ 的正态分布．为确定这种体重是否与七月份相同，七月中旬选取了 25 只雄地鼠的随机样本，并称重得：$\bar{x}=350\text{g}$．假定七月中旬的体重是 $\sigma=100\text{g}$ 的正态分布．则在 0.05 的显著水平下给出：$H_0:\mu=400\text{g}$ 的双侧检验，并以临界值叙述决策规则，给出检验统计量的值以及 P 值，然后说明接受还是拒绝 H_0．

答案：拒绝 H_0，若 $z^*>1.96$ 或 $z^*<-1.96$；$z^*=-2.50$；$P<0.0124$；拒绝 H_0

16.29 已知陆地上的蓝腹蜥蜴的体长总体是 $\mu=60\text{mm}$，$\sigma=15\text{mm}$ 的正态分布．为确定陆地附近岛屿上，同类蜥蜴的总体体长是否有相同大小，从该岛屿选取了 9 只蜥蜴的随机样本，并得 $\bar{x}=65\text{mm}$．假定岛屿上的蜥蜴长度是 $\sigma=15\text{mm}$ 的正态分布，试用临界值决策规则给出 0.05 显著水平下 $H_0:\mu=60\text{mm}$ 的双边检验，并给出检验统计量值及 P 值，然后说明接受还是拒绝 H_0．

答案：拒绝 H_0，若 $z^*>1.96$ 或 $z^*<-1.96$；$z^*=1.00$；$P=0.317$；接受 H_0

16.30 对于习题 16.29 中的蓝腹蜥蜴长度总体，假定用一更大样本（$n=101$）重复研究（$\alpha=0.05$ 下），且同样得到岛屿上的样本值 $\bar{x}=65\text{mm}$．假定岛屿上总体的体长为 $\sigma=15\text{mm}$ 的正态分布，试用临界值决策规则给出 $H_0:\mu=60\text{mm}$ 的双侧检验，并给出检验统计量值及 P 值，然后说明接受还是拒绝 H_0．

答案：拒绝 H_0，若 $z^*>1.96$ 或 $z^*<-1.96$；$z^*=3.35$；$P=0.0008$；拒绝 H_0

16.31 用临界值决策规则给出习题 16.28 中，$n=25$ 的地鼠的样本中 $H_0:\mu=400\text{g}$ 的双侧检验．用 $\bar{X}$ 为检验统计量且 $\alpha=0.01$，给出检验统计量值，然后说明接受还是拒绝 H_0．

答案：拒绝 H_0，若 $\bar{x}^*>[400\text{g}+2.575(\frac{100\text{g}}{\sqrt{25}})=451.5\text{g}]$

或 $\bar{x}^* < [400\text{g} - 2.575(\frac{100\text{g}}{\sqrt{25}}) = 348.5\text{g}]$；$\bar{x}^* = 350\text{g}$；接受 H_0

16.32 现已发现可用一种新的替代牛奶来喂养小牛，并想知道替代奶喂养小牛的第一个月中是否比用牛奶喂养长得更重．已知用牛奶喂养第一个月增加的体重总体是 $\mu = 30\text{lb}$ 和 $\sigma = 12\text{lb}$ 的正态分布，现将36头小牛用替代奶喂养，第一个月平均体重增加为：$\bar{x} = 40\text{lb}$，用 $\bar{X}$ 为检验统计量，$\alpha = 0.05$ 做 $H_0: \mu = 30\text{lb}$的右尾检验．假定新奶喂养下增加的体重是 $\sigma = 12\text{lb}$ 的正态分布，用临界值描述决策规则，并给出检验统计量值，然后说明接受还是拒绝 H_0．

答案：拒绝 H_0，若 $\bar{x}^* > [30\text{lb} + 1.645(\frac{12\text{lb}}{\sqrt{36}}) = 33.29\text{lb}]$；$\bar{x}^* = 40\text{lb}$，拒绝 H_0

β，$1-\beta$，功效曲级及 OC 曲线

16.33 对习题16.32中研究的替代奶问题，若在此替代奶喂养下，小牛体重增加35lb而不是30lb，那么第Ⅱ类错误概率为多少？检验的功效为多少？假定无论是用牛奶还是替代奶喂养的小牛 $\sigma = 12\text{lb}$ 不变．

答案：第Ⅱ类错误概率 $\beta(35) = 0.1949$，检验的功效为 $1 - \beta(35) = 0.8051$

16.34 对习题16.33和16.32中研究的替代奶，当检验中仍用 $\alpha = 0.01$，若平均增长体重为36lb，而不是30lb，那么第Ⅱ类错误概率为多少？检验的功率为多少？同样假定无论是用牛奶还是替代奶喂养的小牛 $\sigma = 12\text{lb}$ 不变．

答案：第Ⅱ类错误概率 $\beta(36) = 0.2514$；检验的功效为 $1 - \beta(36) = 0.7486$

总体均值的假设检验：取自正态分布总体标准差未知的小样本

16.35 一药剂研究员认为对于老年妇女来说，在高原生活会对心率有严重影响，现已知70岁妇女休息时平均心率服从平均每分 $\mu = 74$ 次的正态分布．为探索高原的可能影响，她从居住在9000英尺以上的70岁妇女总体中随机抽取25名的样本．并知平均心率为每分钟 $\bar{x} = 68$ 次及每分钟 $s = 10$ 次，在 $\alpha = 0.05$下给出 $H_0: \mu = 74$ 的双边检验；并给出检验统计量值；说明接受还是拒绝 H_0；给定 H_0 为真下，得到此样本均值的近似概率(界定)．

答案：$t^* = -3.00$；拒绝 H_0；$0.001 < P < 0.01$

16.36 一营养学家开发了一种由葡萄糖和各种离子组成的新型运动饮料，并想知道其是否会改进10公里跑的跑步时间．已知去年参加该市长跑的30岁男子的跑步时间总体为 $\mu = 40\text{min}$ 的正态分布．为确定此运动饮料的影响，他在本次赛跑中随机选取了20名男子为样本，并让他们在下一次10公里赛跑前和中间饮用一定量的该种饮料，结果为：平均跑步时间 $\bar{x} = 38\text{min}$，$s = 5\text{min}$．给出 $\alpha = 0.05$ 下 $H_0: \mu = 40\text{min}$的左尾检验及检验统计量值．请说明接受还是拒绝 H_0；已知 H_0 为真下；得到这个样本均值的近似概率(界定)．

答案：$t^* = -1.789$，拒绝 H_0，$0.025 < P < 0.05$

16.37 习题16.36中的营养学家想对女子检验他的运动饮料．在该市去年10公里赛跑中，30岁女子跑步时间总体为 $\mu = 43\text{min}$ 的正态分布，他随机选取了20名参加这次比赛的女子运动员，在10公里比赛以前和期间让她们饮用一定量的这种饮料．结果为：$\bar{x} = 42\text{min}$，$s = 6\text{min}$．给出 $\alpha = 0.05$ 下 $H_0: \mu = 43\text{min}$ 的左尾检验及检验统计值．请说明接受还是拒绝 H_0；已知 H_0 为真时；得到这个得此样本均值的近似概率(界定)．

答案：$t^* = -0.745$；接受 H_0；$P > 0.10$

正态分布总体的方差和标准差的假设检验

16.38 从习题16.28知八月中旬雄地鼠的体重是 $\mu = 400\text{g}$，$\sigma^2 = 10,000\text{g}^2$ 的正态分布．现感兴趣的是当地鼠生活在更边缘的栖息地，那里的食物的数量与质量随不同位置有很大的变化．地鼠的体重是否也有很大的变化．从更边缘栖息处随机捕捉并称重51只雄地鼠得：$\bar{x} = 350\text{g}$，$s^2 = 11,025\text{g}^2$．作 $H_0: \sigma^2 = 10,000\text{g}^2$ 的右尾检验($\alpha = 0.05$)；给出检验统计量的值；说明接受还是拒绝 H_0；已知 H_0 为真下，得此样本均值的近似概率(界定)．

答案：$\chi^{2*} = 55.125$；接受 H_0，$0.1 < P < 0.5$

16.39 由习题16.29和16.30知陆地的蓝腹蜥蜴的体长总体为 $\mu = 60\text{mm}$ 和 $\sigma^2 = 225\text{mm}^2$ 的正态总体．从习题16.30中来自岛上总体的体长样本($n = 101$)得到：$\bar{x} = 65\text{mm}$ 和 $s^2 = 100\text{mm}^2$．在 $\alpha = 0.05$ 下，作 $H_0: \sigma^2 = 225\text{mm}^2$ 双侧检验；给出检验统计量的值；指出接受还是拒绝 H_0；已知 H_0 为真下，得此样本均值的近似概率(界定)．

答案：$\chi^{2*} = 44.444$；拒绝 H_0，$P < 0.01$

16.40 对习题16.32的替代奶实验，得到牛奶喂养的一个月小牛的增加体重的方差为 $\sigma^2=144\text{lb}^2$. 有明显的证据表明以替代奶喂养的一个月的小牛比以牛奶喂养的小牛体重有较小方差(较小标准差)，重新选取了30只用替代奶喂养一月的小牛为样本，并测其增重，得样本方差 $s^2=81\text{lb}^2$. 在 $\alpha=0.05$ 下，作 $H_0:\sigma=12\text{lb}$ 的左尾检验；给出检验统计量值；请说明接受还是拒绝 H_0；已知 H_0 为真下，得此样本均值的近似概率(界定)

答案：$\chi^{2*}=16.31$；拒绝 H_0，$0.025<P<0.05$

二项总体比率的假设检验

16.41 一鱼类生物学家想知道大湖中小嘴鲈鱼的性别比率. 已知孵化时雄性和雌性数量大致相等，但他认为雄性存活期不如雌性长，并想知道事实上是否成年雄性少于雌性. 他选取了100条成年鱼的随机样本并检验 $H_0:p=0.5$ 对 $H_1:p<0.5$，其中 p 为湖中雄性比率. 给出此检验在 $\alpha=0.01$ 下的临界值决策规则.

答案：拒绝 H_0，若 $z^*<(-z_\alpha=-z_{0.01}=-2.33)$

16.42 就习题16.41中的小嘴鲈鱼，生物学家发现100条鱼的样本中有45条是雄性的($\bar{p}=0.45$). 在 $\alpha=0.01$ 下用临界值决策规则做 $H_0:p=0.5$ 的左尾检验.

答案：$z^*=-1.01$，因为 $z^*>-2.33$，所以接受 H_0

16.43 现想知道潮塘中锯鱼的性别比率，并有理由相信雄性比雌性的少. 随机选取一个潮塘，并检查所有锯鱼，发现有3条为雄性，6条为雌性；在 $\alpha=0.05$ 下，作 $H_0:p=0.5$ 的左尾检验，已知零假设为真下得此样本结果的概率为多少？接受还是拒绝零假设？

答案：$P=P(Y\leqslant 3\mid H_0\text{ 为真})=0.254$；接受 H_0

第十七章　两样本估计和假设检验

17.1　独立样本和成对样本

我们在前三章讨论了单样本参数推断性统计：总体参数的估计（见第十四章和十五章）和参数的假设检验（见第十六章）．现在我们继续讨论两样本参数估计和假设检验．

在一般的两样本问题中，存在两个随机变量 X_1 和 X_2，它们具有类型相同的未知参数（如都有未知均值）．若想知道这些参数有何差别，则可从每个变量（总体）选取观测值样本，然后用它们来估计和检验未知参数的差别．在这样的两样本问题中，样本中每个特定值都用一个双下标符号 x_{ij} 表示，其中 i 表示变量、j 表示样本序列中的特定值．例如 x_{12} 表示变量 X_1 样本中的第二个值．

关于推断性问题，选择合适的两样本方法，必须考虑的许多因素，其中最重要的因素是样本是独立的还是成对的．若从一个样本中选取的观测值，无论怎样也不会影响从另一个样本中选取的观测值的概率，则称样本为**独立样本**（见上册 9.1 节）；若选取的观测值是成对的，本节中的意思是从一个样本中选取的观测值，以某种方式决定了另一个样本中选取的观测值，则称样本为**成对样本**或**相关样本**或**配对样本**．

为理解这些概念，以视力检查中一种用来扩大瞳孔的新型药水的检验问题为例，厂家声称它比标准眼药水见效要快．检验这个说法可用独立样本或成对样本．在独立样本情况中，随机选取两组不同的测试者，其中一组用标准眼药水，另一组用新眼药水；在成对样本情况中，只选一组测试者，并让每个人一只眼睛用标准药水，另一只眼睛用新药水．第一种情况中，由于选取的标准眼药水的观测值怎样也不会影响新眼药水中的观测值，所以样本是独立的；第二种情况中，由于从一个样本得到的观测值与从另一个样本得到的观测值是成对出现的——样本都不是随机选定的，因此样本是成对的．

两种抽样类型各有利弊．通常，两个样本都是随机确定的独立情况是惟一可能的抽样方法．在成对情况中，配对能较好控制外部变量（见上册 3.10 节）．于是在两种眼药水试验中，将两种药水用于同一人可控制样本间关于个体在年龄、健康状况或其他可能影响眼药水反映的外部变量间的随机差别．

我们由独立样本开始研究两样本推断方法．在 17.2 至 17.10 节中，我们将对两总体均值差（$\mu_1-\mu_2$）的估计和假设检验介绍独立样本方法，然后在 17.11 至 17.12 节中介绍成对样本的比较方法．

17.2　两总体均值差（$\mu_1-\mu_2$）的最优估计

现研究两个正态分布总体的关系：第一总体，均值 μ_1 未知，标准差 σ_1 已知；第二总体，均值 μ_2 未知，标准差 σ_2 已知．为估计未知均值差 $\mu_1-\mu_2$，从两个样本总体中分别选取容量为 n_1 和 n_2 的独立随机样本．对这些样本，差的最优估计是什么？

为回答这个问题，我们回到估计量的定义．我们在 14.1 节说过，估计量是给出总体参数 θ 的估计值的任一样本统计量 $\hat{\theta}$，且估计值是由特定样本计算的估计量的数值 $\hat{\theta}^*$．现在这些定义可扩展到两样本情形．

于是我们可以说，若 $\bar{X}_1$ 代表容量为 n_1 的，取自第一总体的随机样本均值，$\bar{X}_2$ 代表容量为 n_2 的，取自第二总体的随机样本均值，且这些样本是独立的，则两样本统计量 $\bar{X}_1-\bar{X}_2$ 是**总体均值差 $\mu_1-\mu_2$ 的一个估计量**．进一步，可由数学方法证明，$\bar{X}_1-\bar{X}_2$ 是 $\mu_1-\mu_2$ 的最优无

偏估计量(见 14.2 节).

我们也可以说在任何选取两个独立样本情形下，计算的样本均值为 $\bar{X}_1=\bar{x}_1$ 和 $\bar{X}_2=\bar{x}_2$，则样本均值差 $\bar{x}_1-\bar{x}_2$ 是**总体均值差 $\mu_1-\mu_2$ 的一个点估计值**.

17.3　均值差的理论抽样分布

我们在 13.6 节说过：均值的理论抽样分布是由对 $\bar{X}$ 的所有可能取值指定的 $f(\bar{x})$ 值组成的概率分布(离散的或连续的).

类似地，在两样本情况，我们可以说：一个离散的或连续的概率分布，若它对两样本统计量 $\bar{X}_1-\bar{X}_2$ 的所有可能值指定了概率或密度，则称这个概率分布为**两均值差的理论抽样分布或均值差的理论抽样分布**. 而且我们可以说(不证明)，若随机变量 X_1 服从均值 μ_1、标准差 σ_1 的正态分布，X_2 服从均值 μ_2，标准差 σ_2 的正态分布，且若对所有容量为 n_1 和 n_2 的，来自两个总体的独立随机样本的可能组合计算 $\bar{X}_1-\bar{X}_2$ 的值，则两均值差的抽样分布服从均值

$$\mu_{\bar{x}_1-\bar{x}_2}=\mu_1-\mu_2 \tag{17.1}$$

方差

$$\sigma^2_{\bar{x}_1-\bar{x}_2}=\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2} \tag{17.2}$$

标准差

$$\sigma_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \tag{17.3}$$

的正态分布. 对一给定的统计量，理论抽样分布的标准差称为该统计量的标准误差(见 13.11 节)，因而(17.3)式称为**均值差的标准误差**.

17.4　均值差($\mu_1-\mu_2$)的置信区间：标准差(σ_1 和 σ_2)已知的正态分布总体的独立样本

如果我们想以已知的肯定程度，将总体参数 θ 置于区间 $a\leqslant\theta\leqslant b$ 中，则需要：(1) 一个样本点估计值 θ^*，(2)一个使点估计值 θ^* 与总体参数 θ 有关的样本统计量，(3)该统计量的抽样分布(见第 14.6 节). 在讨论如何估计两总体均值差 $\mu_1-\mu_2$ 时，我们知道对容量为 n_1 和 n_2 的独立样本，由最优估计量 $\bar{X}_1-\bar{X}_2$ 得到了点估计值 $\bar{x}_1-\bar{x}_2$，并且知道如果 X_1 和 X_2 是正态分布，则 $\bar{X}_1-\bar{X}_2$ 的抽样分布是均值 $\mu_{\bar{x}_1-\bar{x}_2}=\mu_1-\mu_2$，标准差 $\sigma_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ 的正态分布.

这一事实告诉我们，当独立样本是取自均值未知，标准差已知的正态分布总体时，则下列 Z 统计量

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sigma_{\bar{x}_1-\bar{x}_2}}=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_2}+\frac{\sigma_2^2}{n_2}}} \tag{17.4}$$

它可能取的特定值

$$z=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \tag{17.5}$$

满足以上(2)，(3). 它将点估计值 $\bar{x}_1-\bar{x}_2$ 与总体参数 $\mu_1-\mu_2$ 联系起来，由于 $\bar{X}_1-\bar{X}_2$ 服从正态分布，所以这个 Z 统计量的抽样分布是标准正态分布.

于是，用得到 μ 的$(1-\alpha)100\%$置信区间的同样推理方法(见 14.7 至 14.10 节)，我们得

到以下关于 $\mu_1-\mu_2$ 的 $(1-\alpha)100\%$ 置信区间特定值的四种形式:

(1) $(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sigma_{\bar{x}_1-\bar{x}_2}\leqslant\mu_1-\mu_2\leqslant(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sigma_{\bar{x}_1-\bar{x}_2}$;

(2) $L=(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sigma_{\bar{x}_1-\bar{x}_2}$, $U=(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sigma_{\bar{x}_1-\bar{x}_2}$;

(3) $[(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sigma_{\bar{x}_1-\bar{x}_2},(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sigma_{\bar{x}_1-\bar{x}_2}]$;

(4) $(\bar{x}_1-\bar{x}_2)\pm z_{\alpha/2}\sigma_{\bar{x}_1-\bar{x}_2}$.

例 17.1 独立随机样本取自均值 μ_1、μ_2 未知,标准差已知的两个正态分布总体.如果第一总体 $\sigma_1=0.73$,样本值为 $n_1=25$, $\bar{x}_1=6.9$;第二总体 $\sigma_2=0.89$,样本值为 $n_2=20$, $\bar{x}_2=6.7$.那么 $\mu_1-\mu_2$ 的95%置信区间是什么?

解 ☞ 用上面的第四种表示形式,并将(17.3)式的标准误差代入,得

$$(\bar{x}_1-\bar{x}_2)\pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

由表 A.5 得:$z_{0.05/2}=1.96$.于是 95% 置信区间为

$$(6.9-6.7)\pm1.96\sqrt{\frac{(0.73)^2}{25}+\frac{(0.89)^2}{20}}$$

$$0.2\pm0.48$$

17.5 均值差($\mu_1-\mu_2$)的假设检验:标准差(σ_1、σ_2)已知的正态分布总体的独立样本

和单样本假设检验一样,每一个两样本假设检验问题都有一对统计假设:零假设 H_0 和对立假设 H_1.同样也存在双侧和单侧假设检验,而且单侧假设检验又分为右侧检验和左侧检验(见 16.4 和 16.8 节).

在所有两个独立样本假设检验中,零假设为两均值间有某特定差值.如果我们以 δ_0 记此差值(其中 δ 为小写罗马字母 *delta*),则零假设可用符号写成

$$H_0:\mu_1-\mu_2=\delta_0$$

由于 δ_0 可能是任意特定值,所以感兴趣的是更一般的问题:μ_1 和 μ_2 是否有差别.就此问题,零假设为它们并无差别,即 $\delta_0=0$.零假设的其他形式可写成 $H_0:\mu_1-\mu_2=0$ 或 $H_0:\mu_1=\mu_2$.

在两样本假设检验的简要介绍中,我们只考虑 $\delta_0=0$ 的情况,即下列检验给出可能的统计假设:

对于双侧检验,

$$H_0:\mu_1-\mu_2=0,\quad H_1:\mu_1-\mu_2\neq0$$

或

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1\neq\mu_2$$

对于右侧检验,

$$H_0:\mu_1-\mu_2=0,\quad H_1:\mu_1-\mu_2>0$$

或

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1>\mu_2$$

对于左侧检验,

$$H_0:\mu_1-\mu_2=0,\quad H_1:\mu_1-\mu_2<0$$

或

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1<\mu_2$$

我们在16.3节说过,统计假设检验中要得到的是条件概率:已知 H_0 为真,得到与零假设总体参数 θ_0 的差别至少为 $\hat{\theta}^*$ 那样的点估计的概率为多少?这里零假设是 $\mu_1-\mu_2=\delta_0=0$,我们将确定条件概率:

$$P[\text{与 0 至少相差}(\bar{x}_1-\bar{x}_2)\text{ 的点估计值}\mid H_0\text{ 为真}]$$

和单样本假设检验(见16.3节)一样,条件概率的确定需要一个检验统计量:(1)它考虑到样本点估计值与总体参数的零假设值间的比较,(2)它与 H_0 为真假定下的一个已知概率分布有关.对于来自两个标准差已知的正态分布总体的独立样本,可用 Z 统计量[(17.4)式]来检验 $H_0:\mu_1-\mu_2=0$:

$$Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2=0)}{\sigma_{\bar{x}_1-\bar{x}_2}}=\frac{(\bar{X}_1-\bar{X}_2)-0}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \tag{17.6}$$

该统计量通过 $(\bar{x}_1-\bar{x}_2)$ 与0的比较,满足上面的要求(1),又因为如果 X_1 和 X_2 都是正态分布,则在 H_0 为真时,Z 统计量服从标准正态分布,于是它满足要求(2).

对于用这个 Z 统计量的两样本检验,我们给出和单样本检验所用的由 Z 统计量一样的 P 值和临界值决策规则(见16.10节),而对下面例子以及本书中以后的假设检验问题,**我们将同样按16.16节标准检验步骤给出解.**

例17.2 对例17.1的条件,在 $\alpha=0.05$ 下,用临界值决策规则作 $H_0:\mu_1-\mu_2=0$ 的双侧检验.

解 ☞ (1) $H_0:\mu_1-\mu_2=0$, $H_1:\mu_1-\mu_2\neq 0$;

(2) $\alpha=0.05$;

(3) 由于两个总体都是正态分布的,σ_1,σ_2 已知且样本独立,所以我们可以用 Z 统计量[(17.6)式].因为 $\alpha=0.05$,由表A.5得适当值,决策规则为:拒绝 H_0,若 $z^*>(z_{0.05/2}=1.96)$ 或 $z^*<(-z_{0.05/2}=-1.96)$.

(4) Z 统计量的值为

$$z^*=\frac{(6.9-6.7)-0}{\sqrt{\frac{(0.73)^2}{25}+\frac{(0.89)^2}{20}}}=0.810,\text{ 或 }0.81$$

为得到 P 值,用(16.4)式:

$$P=2[P(Z\geqslant(z^*=0.81)\mid H_0\text{ 为真})]$$

它等价于

$$P=2[P(Z>0.81\mid H_0\text{ 为真})]$$

于是,由表A.5得适当值,有

$$P=2(0.5-0.2910)=0.4180$$

(5) 因为 $-1.96<z^*<1.96$,所以在0.05显著水平下,接受 H_0.这一结论可由 $P>0.05$ 确认.

17.6 两均值差的标准误的估计

本章中我们作了不太实际的,即在两总体均值未知时,标准差都是已知的.现在我们开始讨论两个参数集均未知的更一般的情形.在这些条件下,为进行估计和假设检验,我们需要用到由样本计算的均值差的标准误[(17.3)式]的估计,也称为**两均值差的标准误差的估计**(变量记作 $S_{\bar{x}_1-\bar{x}_2}$,计算值记作 $s_{\bar{x}_1-\bar{x}_2}$).依赖于抽样条件和方法要求,这个被估值因而有三种不同形式.

对于有放回的独立抽样,或从无限大或样本容量未知总体抽样,可由下面公式计算 $S_{\bar{x}_1-\bar{x}_2}$:

$$S_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} \tag{17.7}$$

其中 S_1^2 和 S_2^2 是样本方差，n_1 和 n_2 是样本容量.在这些条件下，(17.7)式是 $\sigma_{\bar{x}_1-\bar{x}_2}$ 的最优估计量，此式也可写作

$$S_{\bar{x}_1-\bar{x}_2}=\sqrt{S_{\bar{x}_1}^2+S_{\bar{x}_2}^2} \tag{17.8}$$

其中 $S_{\bar{x}_1}$ 和 $S_{\bar{x}_2}$ 是每一样本计算的均值的标准误.

对于取自有限总体的无放回的独立抽样，其总体容量 N_1 和 N_2 已知，且对每个总体有 $n>0.05N$，将有限总体修正因子加入(17.7)式：

$$S_{\bar{x}_1-\bar{x}_2}=\sqrt{\left[\frac{S_1^2}{n_1}\left(\frac{N_1-n_1}{N_1-1}\right)\right]+\left[\frac{S_2^2}{n_2}\left(\frac{N_2-n_2}{N_2-1}\right)\right]} \tag{17.9}$$

当所用方法(见 17.7 和 17.8 节)需要有总体标准差和方差相等的假定：$\sigma_1=\sigma_2$ 和 $\sigma_1^2=\sigma_2^2$ 时，需要给出 $S_{\bar{x}_1-\bar{x}_2}$ 的第三种形式.上面的假定也称为**方差同质性假定**或**方差齐性假定**.我们将在本章中讨论如何检验此假定(见 17.21 和 17.22 节).但现在，我们可以说，若可假定方差相等($\sigma_1^2=\sigma_2^2=\sigma^2$)，则(17.3)式变为

$$\sigma_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}=\sqrt{\sigma^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \tag{17.10}$$

在上式中，通过用**共同总体方差的合并估计**，即两个样本方差组合在一个公式中，对**共同总体方差** σ^2 进行估计，我们将这个估计量记为 S_p^2：

$$S_p^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} \tag{17.11}$$

将此式代入(17.10)式的 σ^2，并修改为样本形式给出

$$S_{\bar{x}_1-\bar{x}_2}=\sqrt{\left[\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\right]\left(\frac{1}{n_1}+\frac{1}{n_2}\right)} \tag{17.12}$$

或用 S_p^2：

$$S_{\bar{x}_1-\bar{x}_2}=\sqrt{S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}=S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \tag{17.13}$$

这些公式均可假定取特定值

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{\left[\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\right]\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}=s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \tag{17.14}$$

17.7　均值差($\mu_1-\mu_2$)的置信区间：标准差未知，但假定相等($\sigma_1=\sigma_2$)的正态分布总体的独立小样本($n_1<30$ 且 $n_2<30$)

当小样本($n_1<30$ 且 $n_2<30$)是取自均值 μ_1 和 μ_2 未知，标准差 σ_1 和 σ_2 未知的两正态分布总体时，就不能用 Z 统计量[(17.4)式]来计算置信区间，这是因为其中含有未知的总体方差.

然而，如果可以假定方差齐性(见 17.6 节)，则可用 t 分布(见 14.15 节)给出理论上正确的精确解.在这些条件下，所用统计量为

$$T=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{S_{\bar{x}_1-\bar{x}_2}}=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\left[\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\right]\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \quad (17.15)$$

其中,分母为(17.12)式.这个统计量可假定取特定值:

$$t=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s_{\bar{x}_1-\bar{x}_2}}=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\left[\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}\right]\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \quad (17.16)$$

这个统计量满足区间估计的三条标准,这是因为存在:(1) 一个样本点估计值$(\bar{x}_1-\bar{x}_2)$,(2) 一个使点估计值与总体参数$(\mu_1-\mu_2)$有关的样本统计量,(3) 在此问题的条件下,该统计量有一个已知的抽样分布($\nu=n_1+n_2-2$ 个自由度的 t 分布).

因此,用 t 统计量和 14.7 节至 14.10 节的推理方法,可得出两总体的均值差 $\mu_1-\mu_2$ 的 $(1-\alpha)100\%$置信区间.如 17.4 节一样,用这个 t 统计量可以给出四种方式的置信区间:

(1) $(\bar{x}_1-\bar{x}_2)-t_{\alpha/2,\nu}s_{\bar{x}_1-\bar{x}_2}\leqslant\mu_1-\mu_2\leqslant(\bar{x}_1-\bar{x}_2)+t_{\alpha/2,\nu}s_{\bar{x}_1-\bar{x}_2}$;

(2) $L=(\bar{x}_1-\bar{x}_2)-t_{\alpha/2,\nu}s_{\bar{x}_1-\bar{x}_2}$, $\quad U=(\bar{x}_1-\bar{x}_2)+t_{\alpha/2,\nu}s_{\bar{x}_1-\bar{x}_2}$;

(3) $[(\bar{x}_1-\bar{x}_2)-t_{\alpha/2,\nu}s_{\bar{x}_1-\bar{x}_2},(\bar{x}_1-\bar{x}_2)+t_{\alpha/2,\nu}s_{\bar{x}_1-\bar{x}_2}]$;

(4) $(\bar{x}_1-\bar{x}_2)\pm t_{\alpha/2,\nu}s_{\bar{x}_1-\bar{x}_2}$.

例 17.3 为研究睡眠对记忆力的影响,一位心理学家在两种条件下对人群进行实验,内容是有关北极野外生活的纪实电影的细节回忆,这两种条件是:(1) 电影在早七点放映,被测人晚上睡眠正常,第二天晚上给他们 50 个有关电影的多项选择题;(2) 电影早七点放映,被测人白天情况如常,未睡觉,同一天晚上七点给他们 50 个问题.样本是独立的,每组为 15 人,结果为:第(1)组,$n_1=15$, $\bar{x}_1=37.2$ 个正确,$s_1^2=3.33$;第(2)组,$n_2=15$,$\bar{x}_2=35.6$ 个正确,$s_2^2=3.24$.假定两种条件下的总体都是正态分布,且 $\sigma_1^2=\sigma_2^2$,计算 $\mu_1-\mu_2$ 的 95%置信区间.

解 本题可用(17.15)式求解,置信区间为

$$(\bar{x}_1-\bar{x}_2)\pm t_{\alpha/2,\nu}s_{\bar{x}_1-\bar{x}_2}$$

其中$(\bar{x}_1-\bar{x}_2)=37.2-35.6=1.6$.由表 A.6,对于($\nu=n_1+n_2-2=15+15-2=28$)的 t 分布,得 $t_{\alpha/2,\nu}=t_{0.05/2,28}=2.048$.将这些值代入(17.12)式,得

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{\left[\frac{(14\times3.33)+(14\times3.24)}{28}\right]\left(\frac{1}{15}+\frac{1}{15}\right)}=0.6618$$

因此,95%的置信区间为

$$1.6\pm(2.048\times0.6618)$$

$$1.6\pm1.36$$

17.8 均值差$(\mu_1-\mu_2)$的假设检验:标准差未知,但假定相等$(\sigma_1=\sigma_2)$的正态分布总体的独立小样本($n_1<30$ 且 $n_2<30$)

当总体标准差和方差未知,但假定方差齐性时,17.5 节的所有统计假设都可以用.为检验零假设,适合的样本统计量是 $\mu_1-\mu_2=0$ 时的 t 统计量[(17.15)式]:

$$T=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2=0)}{\sqrt{\left[\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}\right]\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \quad (17.17)$$

假定可取特定值:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right]\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{17.18}$$

用这个统计量来确定所需条件概率

$$P[\text{与 0 至少相差}(\bar{x}_1 - \bar{x}_2)\text{ 的点估计值} \mid H_0 \text{ 为真}]$$

由于它满足检验统计量的两个标准:(1) 考虑到点估计值($\bar{x}_1 - \bar{x}_2$)与零假设值($\mu_1 - \mu_2 = 0$)的比较,并以标准误差为单位($s_{\bar{x}_1 - \bar{x}_2}$)表示它们之间距离方式,(2)在 H_0 为真假定下,与已知的概率分布($\nu = n_1 + n_2 - 2$ 个自由度的 t 分布)有关.

只要用合适的自由度,在此可以应用单样本假设检验中的临界值决策规则(见 16.21 节).

例 17.4　对例 17.3 的记忆力研究,在 $\alpha = 0.05$ 下,用临界值决策规则作 $H_0: \mu_1 - \mu_2 = 0$ 的双侧检验.

解 ☞　(1) $H_0: \mu_1 - \mu_2 = 0$,　$H_1: \mu_1 - \mu_2 \neq 0$;

(2) $\alpha = 0.05$;

(3) 因为我们可假定正态分布总体,$\sigma_1^2 = \sigma_2^2$,并且样本是独立的,所以可用(17.17)式.当 $\alpha = 0.05$,($\nu = n_1 + n_2 - 2 = 15 + 15 - 2 = 28$),由表 A.6 得出临界值,双侧检验决策规则为:拒绝 H_0,若 $t^* > (t_{0.05/2,28} = 2.048)$或 $t^* < (-t_{0.05/2,28} = -2.048)$.

(4) t 统计量值为

$$t^* = \frac{1.6 - 0}{0.6618} = 2.4176 \text{ 或 } 2.418$$

$\nu = 28$ 时,近似 P 值在以下范围内[见例 16.6(b)]

$$(2 \times 0.01) < P < (2 \times 0.025) \text{ 或 } 0.02 < P < 0.05$$

(5) 因为 $t^* > 2.048$,所以拒绝 H_0,接受 H_1,这一结论可由 P 值落入范围 $0.02 < P < 0.05$ 来确认.因此,在 0.05 水平下,两样本均值差与零有显著不同,这个结果表明平均地睡眠组[(1)组]分数较高.

17.9　均值差($\mu_1 - \mu_2$)的置信区间:标准差(σ_1 和 σ_2)未知的任何总体分布的独立大样本($n_1 \geq 30$ 且 $n_2 \geq 30$)

当独立的大样本($n_1 \geq 30$ 且 $n_2 \geq 30$)是取自标准差未知的总体,则可根据对总体方差和分布的假定,用 t 或 Z 统计量来确定$(1-\alpha)100\%$的置信区间.

如果在这些条件下,假定方差齐性和正态分布,则在理论上,可用 t 统计量确定精确的置信区间[(17.15)式].然而实际上,如表 A.6 这样的表,一般不能在大样本得到精确 t 值.于是,在两个近似解之间选择一个:(1) 用一个近似 t 值(见 14.23 节),或(2)因为样本容量很大,于是样本标准差是总体标准差的合理估计,这个 Z 统计量的分母是(17.7)式,

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \tag{17.19}$$

可以用它来得到近似置信区间:

如果选取的是大样本,σ_1 和 σ_2 未知,并且总体不能假定为正态分布,则仍可用(17.19)式来确定近似置信区间,因为在这些条件下,中心极限定理(见 13.15 至 13.18 节)保证了 $\bar{X}_1$ 和 $\bar{X}_2$ 的抽样分布是近似正态分布.由此可证明(17.19)式有近似的标准正态分布.

例 17.5　独立随机样本取自于标准差和分布都未知的两个总体.总体(1),均值 μ_1 未知,样本值为 $n_1 = 45$,$\bar{x}_1 = 2.16$,$s_1 = 0.358$.总体(2),均值 μ_2 未知,样本值为 $n_2 = 40$,$\bar{x}_2 = 1.98$,$s_2 = 0.352$.问 $\mu_1 - \mu_2$ 的近似 95%置信区间是什么?

解 ☞　因为是大样本($n_1 \geq 30$ 且 $n_2 \geq 30$),所以可用(17.19)式修正 17.4 节中置信区间,得到近似

95%置信区间为

$$(\bar{x}_1-\bar{x}_2)\pm z_{\alpha/2}s_{\bar{x}_1-\bar{x}_2}$$

其中$(\bar{x}_1-\bar{x}_2)=2.16-1.98=0.18$，对于(17.7)式有

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}=\sqrt{\frac{(0.358)^2}{45}+\frac{(0.352)^2}{40}}=0.0771$$

由表A.5得$z_{\alpha/2}=z_{0.05/2}=1.96$.于是，$\mu_1-\mu_2$的近似95%置信区间为

$$0.18\pm(1.96\times0.0771)$$

$$0.18\pm0.151$$

17.10 均值差($\mu_1-\mu_2$)的假设检验:标准差(σ_1和σ_2)未知的任意分布总体的独立大样本($n_1\geqslant30$且$n_2\geqslant30$)

当独立的大样本($n_1\geqslant30$且$n_2\geqslant30$)是取自标准差未知的总体时，如17.9节中的估计一样，可依据总体方差和分布的假定，用t或Z统计量检验假设.如果假定方差齐性和正态分布，则可用17.9节中$\mu_1-\mu_2=0$时的相同的t和Z统计量.如果不能假定总体为正态分布，则可用$\mu_1-\mu_2=0$时的Z统计量.

例17.6 对例17.5中的条件，在$\alpha=0.05$下，用临界值决策规则作$H_0:\mu_1-\mu_2=0$的双侧检验.

解 ☞ (1) $H_0:\mu_1-\mu_2=0$, $H_1:\mu_1-\mu_2\neq0$;

(2) $\alpha=0.05$;

(3) 在本题条件下，可用$\mu_1-\mu_2=0$时的Z统计量[(17.19)式].当$n_1=45$, $n_2=40$, $\alpha=0.05$时，双侧检验决策规则为:拒绝H_0，若$z^*>(z_{0.05/2}=1.96)$或$z^*<(-z_{0.05/2}=-1.96)$

(4) Z统计量的值为

$$z^*=\frac{(2.16-1.98)-0}{\sqrt{\frac{(0.358)^2}{45}+\frac{(0.352)^2}{40}}}=2.334,\ \text{或}\ 2.33$$

由表A.5，近似P值[(16.4)式]为

$$P\approx2[P(Z\geqslant(z^*=a)\mid H_0\text{为真})]=2[P(Z>2.33)]=2(0.5-0.4901)=0.0198$$

(5) 因为$z^*>1.96$，所以拒绝H_0，接受H_1.这一结论可由近似P值小于0.05确认.于是，在0.05显著水平下，样本差值与零有显著不同，这个结论表明μ_1大于μ_2.

17.11 均值差($\mu_1-\mu_2$)的置信区间:成对样本

回顾17.1节中术语"成对样本"，它的意思是一个样本中的观测值以某种方式决定了从另一个样本选取的观测值.当用这样的成对样本时，需要用到不同于本章目前为止描述的两样本的估计方法.当用成对样本确定两总体均值差($\mu_1-\mu_2$)的置信区间时，适合的方法是**成对样本模型**，其由下列元素组成:(1) n**个差值的一个随机样本**，均值为$\bar{d}$，(2) 样本取自于均值为μ_d的**差值**总体，(3) 样本容量为n的差值样本的**均值差的抽样分布**，均值$\mu_{\bar{d}}=\mu_d$

以X_1和X_2表示所考虑两个总体的随机变量，讨论成对样本模型.两个变量的n对观测值形成成对样本.如果记一对中的每个值为x_{ij}，其中下标i代表变量，j表示样本序列特定值(见17.1节)，则成对样本为

$$(x_{11},x_{21}),(x_{12},x_{22}),\cdots,(x_{1n},x_{2n})$$

如(x_{11},x_{21})表示取自样本中的第一对值，一个值来自于变量X_1，其配对值来自于变量X_2.

下面，对于第j对的每个值，我们讨论x_{1j}和x_{2j}的差，记为d_j，于是有

$$x_{11}-x_{21}=d_1,\ x_{12}-x_{22}=d_2,\ \cdots,\ x_{1n}-x_{2n}=d_n$$

差值样本的均值为

$$\bar{d}=\frac{\sum_{j=1}^{n} d_j}{n} \tag{17.20}$$

标准差为

$$s_d=\sqrt{\frac{\sum_{j=1}^{n}(d_j-\bar{d})^2}{n-1}} \tag{17.21}$$

其可写成下面的计算公式[类似于上册(7.28)式]

$$s_d=\sqrt{\frac{\sum_{j=1}^{n} d_j^2-n\bar{d}^2}{n-1}} \tag{17.22}$$

假定差值样本是取自成对样本差值总体的容量为 n 的随机样本，并有未知均值

$$\mu_d=\frac{\sum_{j=1}^{N} d_j}{N}=\mu_1-\mu_2 \tag{17.23}$$

未知标准差

$$\sigma_d=\sqrt{\frac{\sum_{j=1}^{N}(d_j-\mu_d)^2}{N}} \tag{17.24}$$

还可进一步假定差值总体是正态分布的.

从差值总体取所有容量为 n 的可能随机样本，对每一样本计算均值差 $\bar{d}$，则可得到均值差的理论抽样分布. 因为差值总体是正态分布，所以由随机变量 $\bar{D}$ 表示的这个抽样分布也是正态分布，并且均值为

$$\mu_{\bar{d}}=\mu_d=\mu_1-\mu_2$$

标准差或**均值差的标准误差**(见 13.11 节)为

$$\sigma_{\bar{d}}=\frac{\sigma_d}{\sqrt{n}} \tag{17.25}$$

当 σ_d 未知时，S_d 是 σ_d 的最优估计量，所以**均值差的标准误的估计值**可由上式计算

$$s_{\bar{d}}=\frac{s_d}{\sqrt{n}} \tag{17.26}$$

由此讨论可见，当用成对样本模型时，问题就变为如何用 $\bar{d}$ 估计均值差(μ_d)的单样本问题，而不是如何用 $\bar{x}_1-\bar{x}_2$ 来估计均值差($\mu_1-\mu_2$)的两样本问题. 于是在此单样本形式下，μ_d 的最优估计量为 $\bar{D}$，任意特定样本值 $\bar{D}=\bar{d}$ 是 μ_d 的一个点估计值.

为计算成对样本的总体均值差的$(1-\alpha)100\%$置信区间，我们用 t 统计量：

$$T=\frac{\bar{D}-\mu_d}{S_{\bar{d}}} \tag{17.27}$$

它可取特定值

$$t=\frac{\bar{d}-\mu_d}{s_{\bar{d}}} \tag{17.28}$$

可证明 t 统计量是 $\nu=n-1$ 个自由度的 t 分布，其中 $n=$(总对数). 于是，精确解是 14.21 节描述的单样本 t 解，并用 $\bar{d}$ 代替 $\bar{x}$，μ_d 代替 μ，$s_{\bar{d}}$ 代替 $s_{\bar{x}}$，因此，μ_d 的精确的$(1-\alpha)100\%$置信区间可由下列形式之一表示：

(1) $\bar{d}-t_{\alpha/2,\nu}s_{\bar{d}}\leqslant\mu_d\leqslant\bar{d}+t_{\alpha/2,\nu}s_{\bar{d}}$；

(2) $L=\bar{d}-t_{\alpha/2,\nu}s_{\bar{d}}$，　$U=\bar{d}+t_{\alpha/2,\nu}s_{\bar{d}}$；

(3) $[\bar{d} - t_{\alpha/2,\nu}s_{\bar{d}},\ \bar{d} + t_{\alpha/2,\nu}s_{\bar{d}}]$;

(4) $\bar{d} \pm t_{\alpha/2,\nu}s_{\bar{d}}$.

对任何样本容量为 n 该 t 解是精确解,但若 $n \geqslant 30$,则 Z 统计量

$$Z = \frac{\bar{D} - \mu_d}{S_{\bar{d}}} \tag{17.29}$$

有近似的标准正态分布,这是因为 $S_{\bar{d}}$ 是 $\sigma_{\bar{d}}$ 的合理估计量.于是,当 $n \geqslant 30$ 时,得到两个总体均值差的置信区间的近似解是单样本 Z 解(见 14.10 和 14.23 节),用 $\bar{d}$ 代替 $\bar{x}$,μ_d 代替 μ,$s_{\bar{d}}$ 代替 $s_{\bar{x}}$. μ_d 的近似 $(1-\alpha)100\%$ 置信区间可由下列形式之一表示:

(1) $\bar{d} - z_{\alpha/2}s_{\bar{d}} \leqslant \mu_d \leqslant \bar{d} + z_{\alpha/2}s_{\bar{d}}$;

(2) $L = \bar{d} - z_{\alpha/2}s_{\bar{d}},\ U = \bar{d} + z_{\alpha/2}s_{\bar{d}}$;

(3) $[\bar{d} - z_{\alpha/2}s_{\bar{d}},\ \bar{d} + z_{\alpha/2}s_{\bar{d}}]$;

(4) $\bar{d} \pm z_{\alpha/2}s_{\bar{d}}$.

例 17.7　患有高血压的 10 名男患者正在进行一种锻炼与饮食的系统调养,以期降低血压.表 17.1 给出每个病人训练开始时(x_{1j})和六个月后的(x_{2j})舒张压(单位:mmHg).由此数据集,试确定 μ_d(开始时与六个月后的差的均值)的 99% 置信区间,假定差值总体服从正态分布.

表 17.1

病人 j	起始 x_{1j}(mmHg)	六个后 x_{2j}(mmHg)
1	141	142
2	169	165
3	158	150
4	180	176
5	147	143
6	160	157
7	175	170
8	163	157
9	148	143
10	163	162

表 17.2

病人 j	起始 x_{1j}(mmHg)	六个月后 x_{2j}(mmHg)	d_j(mmHg)	d_j^2(mmHg)^2
1	141	142	−1	1
2	169	165	4	16
3	158	150	8	64
4	180	176	4	16
5	147	143	4	16
6	160	157	3	9
7	175	170	5	25
8	163	157	6	36
9	148	143	5	25
10	163	162	1	1
$\sum$			39mmHg	209(mmHg)^2

$$\bar{d} = \frac{\sum d_j}{n} = \frac{39\text{mm}}{10} = 3.9\text{mmHg}$$

$$s_d = \sqrt{\frac{\sum d_j^2 - n\bar{d}^2}{n-1}} = \sqrt{\frac{209(\text{mmHg})^2 - [10 \times (3.9\text{mmHg})^2]}{10-1}} = 2.5144\text{mmHg}$$

$$s_{\bar{d}} = \frac{s_d}{\sqrt{n}} = \frac{2.5144\text{mmHg}}{\sqrt{10}} = 0.7951\text{mmHg}$$

解　精确解为

$$\bar{d} \pm t_{\alpha/2,\nu} s_{\bar{d}}$$

为确定 $\bar{d}$ 和 $s_{\bar{d}}$，我们在表 17.1 中加上两列：差值列($d_i = x_{1i} - x_{2i}$)和差值平方列(d_i^2). 添加这些列后，以及用(17.20)、(17.22)和(17.26)式计算 $\bar{d}$，s_d 和 $s_{\bar{d}}$ 的结果均在表 17.2 列出.

由表 A.6，当 $\nu = 10 - 1 = 9$，$t_{0.01/2,9} = 3.250$ 时，99%的置信区间为

$$3.9\text{mmHg} \pm (3.250 \times 0.7951\text{mmHg})$$

$$3.9\text{mmHg} \pm 2.58\text{mmHg}$$

17.12　均值差($\mu_1 - \mu_2$)的假设检验：成对样本

同单样本假设检验和独立样本假设检验一样，成对样本的每个两样本假设检验问题都有一对统计假设：H_0 和 H_1，而且和以前一样，也存在双侧和单侧检验. 在这些检验中，零假设是在两均值 $\mu_1 - \mu_2$ 间存在特定差值 δ_0(见 17.5 节). 对于成对样本，$\mu_1 - \mu_2 = \mu_d$[(17.23)式]，于是 H_0 可记为 $H_0: \mu_1 - \mu_2 = \delta_0$ 或 $H_0: \mu_d = \delta_0$.

同样，和 17.5 节一样，我们只考虑 $\delta_0 = 0$ 的情况，检验的可能统计假设为：对于双侧检验

$$H_0: \mu_1 - \mu_2 = 0, \quad H_1: \mu_1 - \mu_2 \neq 0$$

或

$$H_0: \mu_d = 0, \quad H_1: \mu_d \neq 0$$

对于右侧检验

$$H_0: \mu_1 - \mu_2 = 0, \quad H_1: \mu_1 - \mu_2 > 0$$

或

$$H_0: \mu_d = 0, \quad H_1: \mu_d > 0$$

对于左侧检验

$$H_0: \mu_1 - \mu_2 = 0, \quad H_1: \mu_1 - \mu_2 < 0$$

或

$$H_0: \mu_d = 0, \quad H_1: \mu_d < 0$$

当用成对样本检验 $H_0: \mu_d = 0$ 时，如果可以假定差值总体是正态分布的，则当 $\mu_d = 0$ 时，由 t 统计量[(17.27)式]得精确解为

$$T = \frac{\bar{D} - (\mu_d - 0)}{S_{\bar{d}}} \tag{17.30}$$

可取特定值为

$$t = \frac{\bar{d} - (\mu_d = 0)}{s_{\bar{d}}} \tag{17.31}$$

可用这个统计量得到所需条件概率(见 17.5 节)：

$$P(\text{与 0 至少相差 } \bar{d} \text{ 的点估计值} \mid H_0 \text{ 为真})$$

因为它满足检验统计量的两个标准(见 16.3 节)：(1)它比较 $\bar{d}$ 和 $\mu_d = 0$，并以标准误差为单位表示它们间的距离，(2) 在已知 H_0 为真假定下，有已知的概率分布($\nu = n - 1$ 个自由度的 t 分布，其中 n = 总对数). 对此假设检验，以合适的自由度，可用单样本假设检验中的 P 值和决

策规则(见 16.20 和 16.21 节).

当 $n \geqslant 30$ 时,在 $\mu_d=0$ 下,用 Z 统计量[(17.29)式]可得到成对样本假设检验的近似解.

$$Z=\frac{\bar{D}-(\mu_d=0)}{S_{\bar{d}}} \tag{17.32}$$

它可取特定值

$$z=\frac{\bar{d}-(\mu_d=0)}{s_{\bar{d}}} \tag{17.33}$$

这个 Z 统计量也满足检验统计量的两个标准:(1) 它比较 $\bar{d}$ 与零,(2) 在已知 H_0 为真假定下,有近似的标准正态分布.对此假设检验,可用单样本假设检验中的 P 值和决策规则(见 16.6 至 16.10 节).

例 17.8 对例 17.7 中的血压研究,在 $\alpha=0.01$ 下,用临界值决策规则作 $H_0:\mu_d=0$ 的右尾检验.

解 (1) $H_0:\mu_d=0 \quad H_1:\mu_d>0$.

(2) $\alpha=0.01$.

(3) 由于样本是成对的,并且假定差值(d_i)总体是正态分布,所以可用 t 统计量[(17.30)式].于是由 $\mu_d=0$,$\alpha=0.01$,$\nu=n-1=10-1=9$ 以及由表 A.6 得的适当值,右侧决策规则为:拒绝 H_0,若 $t^*>(t_{0.01,9}=2.821)$.

(4) t 统计量的值为

$$t^*=\frac{3.9\text{mmHg}}{0.7951\text{mmHg}}=4.9050 \text{ 或 } 4.905$$

$\nu=9$ 时的 P 值在 $P<0.0005$ 的范围中.

(5) 因为 $t^*>2.821$,所以拒绝 H_0,接受 H_1.这一结果可由 P 值小于 0.01 确认.于是可以说,在 0.01 显著水平下,没有明显的理由表明,平均来说在锻炼与饮食系统调养六个月后舒张压下降了.

17.13 均值的两样本参数估计和假设检验的假定

我们在 16.24 节讨论了单样本参数估计和假设检验的假定.那时指出这些假定是数学模型的定义性质,这些模型形成了每种方法的基础:被研究总体以及从中抽取的样本所需的特征.这些方法之所以称为参数方法,是因为它们对参数进行估计和假设检验,并且许多假设都是关于参数的.在 17.4 至 17.12 节中描述的两样本均值差的估计和假设检验的方法,由于同样原因也称为参数方法,而且已经看到,它们的数学模型与单样本模型有许多相同的假定.另外,还有两样本问题特有的假定.

对 17.4 和 17.5 节中的模型,假定所研究总体是正态分布的间隔水平和比例水平测量值,而总体标准差已知.假定样本为任意样本容量的独立(见 17.1 节)随机样本,其他适用于 16.5 节模型的抽样假定也适用于此,以及本章所讨论的其他模型.

对 17.7 和 17.8 节中的模型,同样假定所研究总体是正态分布的间隔水平和比例水平测量值,虽然总体标准差未知,但假定方差相等:方差齐性(见 17.6 节).同样假定样本为任意样本容量的独立随机样本.

17.9 和 17.10 节中的模型都假定,独立的大样本($n_1 \geqslant 30$ 且 $n_2 \geqslant 30$)是取自以间隔水平或比例水平测量值的总体,且总体标准差未知.如果假定方差齐性和正态分布,则可用精确的或近似的 t 方法或近似 Z 方法进行估计和假设检验.如果在这些条件下不能假定正态性,则有基于中心极限定理的近似方法.

17.11 和 17.12 节的模型与以前的两样本模型有很多不同的假定.现在由于样本是成对的,不能再假定为独立随机样本.但是,如果将每一对变为一个差值,并假定差值的样本结果是来自于差值的正态分布总体的随机样本,则可用精确的单样本 t 解来进行估计和假设检验.反之,如果(差值个数)$\geqslant 30$,则可假定 $s_{\bar{d}}$ 是 $\sigma_{\bar{d}}$ 的一个合理估计量,于是可用近似的 Z 解.

17.14 如果违背假定

在均值的单样本参数估计和假设检验假定的讨论中(见 16.24 节),我们指出:除了抽样要求的多样性,还有三个基本考虑:(1) 总体是否可假定为正态分布,(2) σ 是否已知,(3) 样本容量.现在,在更为复杂的双样本情形中,有四个基本考虑:(1) 两个总体是否都可假定为正态分布,(2) σ_1 和 σ_2 是否都已知,(3) 是否可假定 $\sigma_1=\sigma_2$,由此 $\sigma_1^2=\sigma_2^2$(方差齐性),(4) 两个样本容量.

和单样本问题一样,两样本问题的第一步是分析样本数据满足假定的情况.16.25 节给出了检查正态分布假定的方法.我们将在 17.21 节给出如何检验方差的齐性.但现在可以说独立样本 t 方法在假定违背时比齐次检验更稳健,尤其是在样本容量相等 $n_1=n_2$ 情况下,齐次假定不适用于成对样本检验.

如果已知与假定存在严重的偏离,则如单样本方法时一样,下一步是试图数据的变换(见 16.25 节).同样不存在固定的变换法则,但对不同背离有一些建议.比如,如果分布是正偏的,对数变换可使分布更对称,且使方差相等.一般地,如果标准差与均值成比例,则可用对数变换;如果方差与均值成比例,则可试试平方根变换.对计数数据,平方根变换可改善正态性和方差齐性.如果分布是负偏的,或者方差随均值增大而减小,则可试试将数据平方.

和单样本方法一样,如果变换成功,则可对变换后的数据进行估计和假设检验,并在结果中指明变换方法.

如果原始数据严重违反所有可用参数方法的假定,而且找不到满足假定的变换,则不可能确定 $\mu_1-\mu_2$ 的 $(1-\alpha)100\%$ 置信区间,或检验 $H_0:\mu_1-\mu_2=0$.然而,还是和单样本过程一样,可用非参数方法分析数据的其他方面(见第 20 章).

17.15 独立样本和成对样本方法在精确性和功效方面的比较

在许多情形中,为解决估计和假设检验问题,总可在成对样本方法和独立样本方法之间作一选择.一般来说,如果可得到正确的配对(见 17.1 节),并且所用的对数 n 等于每一个独立样本容量($n_1=n_2=n$).在此条件下,成对样本方法更好些.这是因为,对同样的水平 α,成对值的置信区间将更精确(更窄,见 14.11 节),而且成对样本假设检验比独立样本检验功效更强(更可能拒绝错误的零假设,见 16.18 节).

17.16 F 统计量

假设现研究两个正态分布总体间的关系.第一个总体,参数 μ_1,σ_1^2 和 σ_1 未知,第二个总体,参数 μ_2,σ_2^2 和 σ_2 未知.为进行关于这些参数间比较的统计推断,从两个总体分别取容量为 n_1 和 n_2 的独立随机样本.由本章可知,若进行两总体均值比较的推断,需要用样本差值 $\bar{x}_1-\bar{x}_2$ 和 t 或 Z 方法估计和检验它们的差值 $\mu_1-\mu_2$.但在此问题条件下,用什么样的统计方法来进行两个总体方差 σ_1^2 和 σ_2^2 比较的推断呢?

在这些条件下,比较总体方差不能用方差的差 $\sigma_1^2-\sigma_2^2$ 来进行统计推断,而是用**方差比** σ_1^2/σ_2^2.由于 S_1^2 是 σ_1^2 的最优估计量,S_2^2 是 σ_2^2 的最优估计量(见 14.2 节),所以 S_1^2/S_2^2 是 σ_1^2/σ_2^2 的最优估计量.因此任意特定的样本方差比 $S_1^2/S_2^2=s_1^2/s_2^2$ 是总体比 σ_1^2/σ_2^2 的一个点估计值.

为对 σ_1^2/σ_2^2 进行估计和假设检验,我们需要将任意特定点估计值 s_1^2/s_2^2 与 σ_1^2/σ_2^2 联系起来的样本统计量,并且有已知的抽样分布.鉴于此,我们将用 F 统计量

$$F=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}=\frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} \tag{17.34}$$

它可取特定值

$$f=\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}=\frac{s_1^2\sigma_2^2}{s_2^2\sigma_1^2} \tag{17.35}$$

对于我们提出的抽样条件,这个统计量有一个已知的抽样分布,称为 **F 分布**(见 17.17 节).

在 F 统计量中,哪个总体及样本记为 1 并置于分子中是任意的.但当较大的样本方差记为 S_1^2 作为分子时,用 F 分布临界值表(见 17.19 节)比较容易.某些统计著作中推荐这种方法.在本书中,我们所用方法不需将较大的样本方差置于分子中.

17.17　F 分布

由抽样理论知:

给定两个独立的 χ^2 随机变量(见 13.2 和 15.2 节):X_1^2,X_2^2,则这两个连续变量分别被其自由度 ν_1, ν_2 除后的比值是一个连续的 F **随机变量**.

$$F=\frac{X_1^2/\nu_1}{X_2^2/\nu_2} \tag{17.36}$$

它是有两种自由度形式的 F 分布:ν_1 是 F 比分子的自由度,ν_2 是 F 比分母的自由度.

这个统计量可取特定值:

$$f=\frac{\chi_1^2/\nu_1}{\chi_2^2/\nu_2} \tag{17.37}$$

再者,对于 17.16 节中介绍的抽样条件,来自正态分布的独立样本,下列变量是独立的 χ^2 变量(见 15.2 节):

$$X_1^2=\frac{(n_1-1)S_1^2}{\sigma_1^2},\quad \nu_1=n_1-1 \text{ 自由度}$$

$$X_2^2=\frac{(n_2-1)S_2^2}{\sigma_2^2},\quad \nu_2=n_2-1 \text{ 自由度}$$

于是,(17.34)式和(17.36)式的 F 统计量.

$$F=\frac{X_1^2/\nu_1}{X_2^2/\nu_2}=\frac{\dfrac{(n_1-1)S_1^2}{\sigma_1^2(n_1-1)}}{\dfrac{(n_2-1)S_2^2}{\sigma_2^2(n_2-1)}}=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}=\frac{S_1^2/\sigma_2^2}{S_2^2/\sigma_1^2}$$

有一个分子自由度为 $\nu_1=n_1-1$,分母自由度为 $\nu_2=n_2-1$ 的 F 分布.由此统计量可看出,为什么 F 分布也称为**方差比分布**.

和以前所有的概率分布一样,F 分布也由一个特定的惟一的概率函数所定义.不加推导,我们给出 F 分布的连续概率密度函数为

$$f(f)=cf^{\left(\frac{\nu_1}{2}\right)-1}\left(1+\frac{\nu_1 f}{\nu_2}\right)^{-\left(\frac{\nu_1+\nu_2}{2}\right)} \tag{17.38}$$

其中 $f=\dfrac{\chi_1^2/\nu_1}{\chi_2^2/\nu_2}=\dfrac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$

$\nu_1=F$ 统计量分子自由度

$\nu_2=F$ 统计量分母自由度

$c=$常数,使 F 分布下总面积为 1 的 ν_1 和 ν_2 的函数.

t 分布和 χ^2 分布仅有一个自由度参数 ν(见 14.15 和 15.2 节)，而 F 分布有两个自由度参数 ν_1 和 ν_2. 以分子自由度叙述在先，这个表示顺序是必须约定的，以后我们会给出理由.

和以前概率分布一样，F 分布不仅仅是一个分布，而是一个连续概率分布族，每个正整数值自由度的组合都是一个 F 分布. 这无限多个分布统称为 F 分布. 图 17-1 给出两个这样的分布(一个是 $\nu_1=2$, $\nu_2=2$. 另一个是 $\nu_1=9$, $\nu_2=9$).

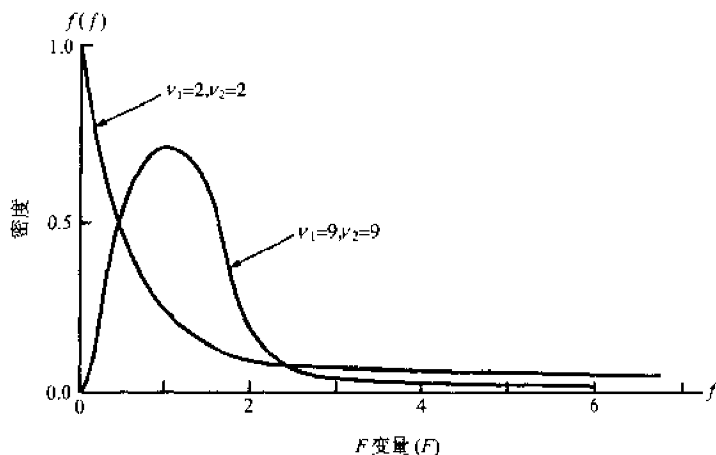


图 17-1

因为变量 F 是两个独立的 χ^2 变量被本身各自的自由度相除后的比，所以 F 分布(见图 17-1)很像 χ^2 分布(见图 15-1)便不足为奇. 两个分布都是连续的、只取正值的不对称(正偏)分布，而且和 χ^2 分布随着 ν 的增加越来越对称一样，F 分布随 ν_1 和 ν_2 的增加也越来越对称. F 分布的均值为

$$\mu = \frac{\nu_2}{\nu_2 - 2}, \quad \nu_2 > 2 \tag{17.39}$$

随着 ν_2 的增加它接近于 1. F 分布的方差为

$$\sigma^2 = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}, \quad \nu_2 > 4 \tag{17.40}$$

标准差为 $\sqrt{\sigma^2}$.

F 分布是以英国统计学家 Ronaid Aylmer Fisher 的名字(1890-1962)命名的，作为**方差分析**(见第十八章)研究中的一部分，他得到并列表给出 F 分布. 原来这个分布称为方差比分布，后来 Fisher 的学生为纪念他而改名为 F 分布. 除了 F 分布和方差分析以外，Fisher 还为统计学做出了许多重要贡献，其中包括实验设计，相关(见第十九章)和假设检验理论等基础工作.

17.18　F 分布的临界值

我们在 15.3 节说过，对于随机变量

$$X^2 = \frac{(n-1)S^2}{\sigma^2}$$

当样本来自于正态分布时，它有 $\nu = n-1$ 的 χ^2 分布，则[(15.4)式]

$$P(X^2 > \chi^2_{\alpha,\nu}) = \alpha$$

类似地，我们说对于两个正态总体的独立抽样，随机变量 F[(17.34)和(17.36)式]有 $\nu_1 = n_1 - 1$ 和 $\nu_2 = n_2 - 1$ 的 F 分布，

$$P(F > f_{\alpha, \nu_1, \nu_2}) = \alpha \tag{17.41}$$

其中 f_{α, ν_1, ν_2} 表示右侧为 α,左侧为 $1-\alpha$ 的自由度为 ν_1 和 ν_2 的 F 分布的临界值,面积与临界值间的关系如图 17-2 所示.

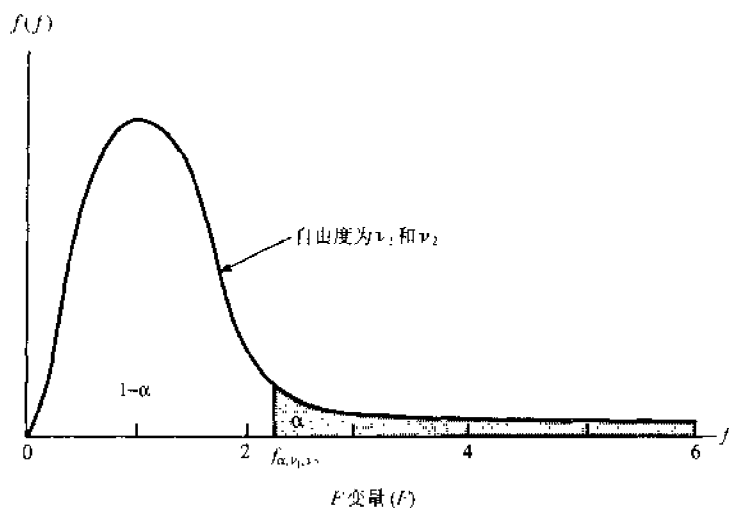


图 17-2

对于有不对称分布的随机变量 X^2,我们在 15.3 节进一步指出[(15.5)式]

$$P(X^2 < \chi^2_{1-\alpha, \nu}) = \alpha$$

类似地,对于随机变量 F,我们给出

$$P(F < f_{1-\alpha, \nu_1, \nu_2}) = \alpha \tag{17.42}$$

其中 $f_{1-\alpha, \nu_1, \nu_2}$ 表示右侧为 $1-\alpha$,左侧为 α 的自由度为 ν_1 和 ν_2 的 F 分布的临界值,面积与临界值间的关系如图 17-3 所示.

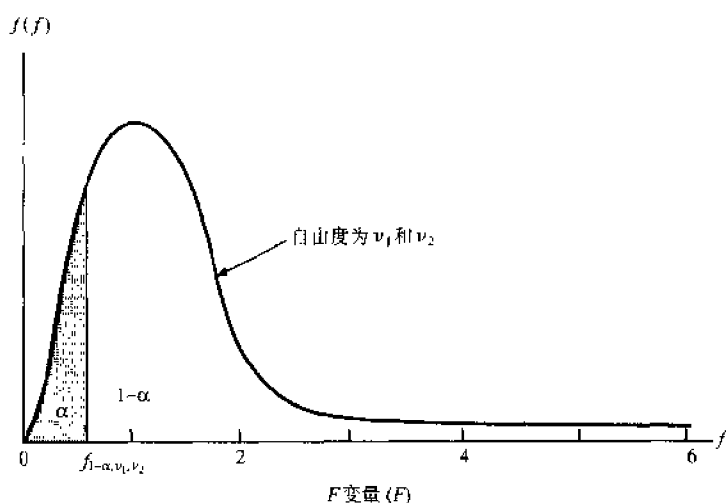


图 17-3

$f_{1-\alpha, \nu_1, \nu_2}$ 和 f_{α, ν_2, ν_1} 之间具有倒数关系,即

$$f_{1-\alpha,\nu_1,\nu_2} = \frac{1}{f_{\alpha,\nu_2,\nu_1}} \quad (17.43)$$

也称为 $f_{1-\alpha,\nu_1,\nu_2}$ 的**倒数法则**或**倒数性质**. 注意到 f_α 的倒数值也改变了自由度的顺序:(ν_1,ν_2)变为(ν_2,ν_1). 当我们用的表 A.8:**F 分布的临界值**(见 17.19 节)时,这个变换很重要.

最后,我们在 15.3 节指出,对于随机变量 X^2

$$P(\chi^2_{1-\alpha/2,\nu} \leqslant X^2 \leqslant \chi^2_{\alpha/2,\nu}) = 1-\alpha$$

类似地,对于随机变量 F,

$$P(f_{1-\alpha/2,\nu_1,\nu_2} \leqslant F \leqslant f_{\alpha/2,\nu_1,\nu_2}) = 1-\alpha \quad (17.44)$$

其中 $f_{\alpha/2,\nu_1,\nu_2}$ 表示自由度为 ν_1 和 ν_2 的 F 分布的右侧面积为 $\alpha/2$ 的临界值,$f_{1-\alpha/2,\nu_1,\nu_2}$ 记为同一 F 分布的左侧面积为 $\alpha/2$ 的临界值. 图 17-4 给出了临界值与面积间的关系. 倒数规则[(17.43)式]也适用于双侧值.

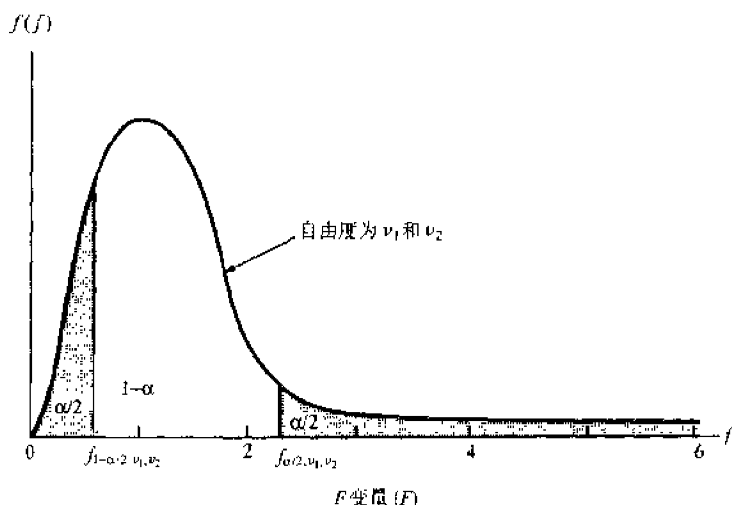


图 17-4

17.19　表 A.8:F 分布的临界值

表 A.8 给出了部分 F 分布的五个 α 水平(0.100,0.050,0.025,0.010,0.005)的 f_{α,ν_1,ν_2} 临界值. F 分布是由分子自由度(顶部一行:$\nu_1 = 1, \cdots, 10, 12, 15, 20, 24, 30, 40, 60, 120, \infty$)和分母自由度(左侧一列:$\nu_2 = 1,\cdots,30,40,60,120,\infty$)来标识的. 表 A.8 仅给出了 f_α 的值,用倒数规则[(17.43)式]可得到 $f_{1-\alpha}$ 的值.

例 17.9　用表 A.8 求 $f_{\alpha,\nu_1,\nu_2} = f_{0.050,7,8}$ 和 $f_{1-\alpha,\nu_1,\nu_2} = f_{1-0.050,7,8}$ 的值.

解　表 17.3 为表 A.8 的一部分,其中圈画出了本题解. 为得到 $f_{0.050,7,8}$ 的值,先找到 $\nu_1 = 7$ 列,然后由此列找到标记 $\nu_2 = 8$ 的第五行(每行为一个 α 水平),其中在 $\alpha = 0.050$ 处,我们找到并圈出 3.50. 因此 $f_{0.050,7,8} = 3.50$ 是 $\nu_1 = 7$, $\nu_2 = 8$ 的 F 分布右尾面积为 $\alpha = 0.050$ 的临界值.

为得到 $f_{1-0.050,7,8}$ 的值,用(17.43)式,有

$$f_{1-0.050,7,8} = \frac{1}{f_{0.050,8,7}}$$

为得到 $f_{0.050,8,7}$ 的值,我们首先习惯上将 ν_2 放在前边(见 17.17 节),将它看成是表 A.8 中的"ν_1". 于是,在表 17.3 中找到 $\nu_1 = 8$ 列,然后沿着此列找到 $\nu_2 = 7$ 行,其中在 $\alpha = 0.050$ 处圈出值 3.73. 因此,

表 17.3

ν_2	α	ν_1								
		1	2	3	4	5	6	7	8	9
7	0.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	0.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	0.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
	0.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	0.005	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51
8	0.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	0.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	0.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	0.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	0.005	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34

$$f_{1-0.050,7,8}=\frac{1}{3.73}=0.2681\text{ 或 }0.27$$

于是,$f_{1-0.050,7,8}=0.27$ 是 $\nu_1=7$, $\nu_2=8$ 的 F 分布左尾 $\alpha=0.050$ 的临界值.

17.20 方差比(σ_1^2/σ_2^2)的置信区间:参数(σ_1^2, σ_1, μ_1 和 σ_2^2, σ_2, μ_2)未知的正态分布总体的独立样本

由取自正态分布总体的独立随机样本可知,随机变量 F[(17.34)和(17.36)式]有 $\nu_1=n_1-1$, $\nu_2=n_2-1$ 的 F 分布(见 17.17 节). 由(17.44)式知,对于这个 F 分布有

$$P(f_{1-\alpha/2,\nu_1,\nu_2}\leqslant F\leqslant f_{\alpha/2,\nu_1,\nu_2})=1-\alpha$$

由这一概率表述,我们可以得到方差比 σ_1^2/σ_2^2 的置信区间. 首先,将(17.34)式代入(17.44)式得

$$P\left(f_{1-\alpha/2,\nu_1,\nu_2}\leqslant\frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2}\leqslant f_{\alpha/2,\nu_1,\nu_2}\right)=1-\alpha$$

每项乘以 S_2^2/S_1^2 得

$$P\left[\frac{S_2^2}{S_1^2}(f_{1-\alpha/2,\nu_1,\nu_2})\leqslant\frac{\sigma_2^2}{\sigma_1^2}\leqslant\frac{S_2^2}{S_1^2}(f_{\alpha/2,\nu_1,\nu_2})\right]=1-\alpha$$

每项取倒数得

$$P\left[\frac{S_1^2}{S_2^2}\left(\frac{1}{f_{1-\alpha/2,\nu_1,\nu_2}}\right)\geqslant\frac{\sigma_1^2}{\sigma_2^2}\geqslant\frac{S_1^2}{S_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right)\right]=1-\alpha$$

可写成

$$P\left[\frac{S_1^2}{S_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right)\leqslant\frac{\sigma_1^2}{\sigma_2^2}\leqslant\frac{S_1^2}{S_2^2}\left(\frac{1}{f_{1-\alpha/2,\nu_1,\nu_2}}\right)\right]=1-\alpha$$

最后,由倒数法则[(17.43)式]得

$$P\left[\frac{S_1^2}{S_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right)\leqslant\frac{\sigma_1^2}{\sigma_2^2}\leqslant\frac{S_1^2}{S_2^2}\left(\frac{1}{\dfrac{1}{f_{\alpha/2,\nu_2,\nu_1}}}\right)\right]=1-\alpha$$

于是有

$$P\left[\frac{S_1^2}{S_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right)\leqslant\frac{\sigma_1^2}{\sigma_2^2}\leqslant\frac{S_1^2}{S_2^2}(f_{\alpha/2,\nu_2,\nu_1})\right]=1-\alpha \qquad (17.45)$$

在从正态分布总体独立随机抽样的条件下，F 统计量[(17.34)和(17.36)式]符合区间估计的标准(见14.6节)，这是因为存在：(1) 总体参数(σ_1^2/σ_2^2)的样本点估计量(s_1^2/s_2^2)；(2) 使点估计量与总体参数有关的样本统计量(F)；(3)在这些条件下，已知的抽样分布($\nu_1=n_1-1$，$\nu_2=n_2-1$ 的 F 分布)．

由(17.45)式知，如果对每个样本计算随机区间

$$\frac{S_1^2}{S_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right)\leqslant\frac{\sigma_1^2}{\sigma_2^2}\leqslant\frac{S_1^2}{S_2^2}(f_{\alpha/2,\nu_2,\nu_1})$$

的特定值，则特定区间：

$$\frac{s_1^2}{s_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right)\leqslant\frac{\sigma_1^2}{\sigma_2^2}\leqslant\frac{s_1^2}{s_2^2}(f_{\alpha/2,\nu_2,\nu_1})$$

的$(1-\alpha)100\%$含有 σ_1^2/σ_2^2．因此，任何一个这样的区间都是 σ_1^2/σ_2^2 的$(1-\alpha)100\%$置信区间．和单个总体方差的$(1-\alpha)100\%$置信区间(见15.6节)一样，总体方差比的置信区间也可写成三种形式：

(1) $\frac{s_1^2}{s_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right)\leqslant\frac{\sigma_1^2}{\sigma_2^2}\leqslant\frac{s_1^2}{s_2^2}(f_{\alpha/2,\nu_2,\nu_1})$

(2) $L=\frac{s_1^2}{s_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right)$，$U=\frac{s_1^2}{s_2^2}(f_{\alpha/2,\nu_2,\nu_1})$

(3) $\left[\frac{s_1^2}{s_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right),\ \frac{s_1^2}{s_2^2}(f_{\alpha/2,\nu_2,\nu_1})\right]$

例 17.10　例17.3中的记忆力研究，给出 σ_1^2/σ_2^2 的近似95%置信区间．

解　在例17.3中，独立样本($n_1=15$，$s_1^2=3.33$ 和 $n_2=15$，$s_2^2=3.24$)取自两个正态分布总体．于是，可以用$(1-\alpha)100\%$置信区间：

$$\left[\frac{s_1^2}{s_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right),\ \frac{s_1^2}{s_2^2}(f_{\alpha/2,\nu_2,\nu_1})\right]$$

其中$\frac{s_1^2}{s_2^2}=\frac{3.33}{3.24}$．然而到表A.8中找临界值，我们发现两个问题：(1) 表中没有 $\alpha/2$ 的临界值，(2) 表中也没有 $\nu_1=n_1-1=14$，$\nu_2=n_2-1=14$ 的 F 分布的临界值．第一个问题很好解决，只要将 $f_{\alpha/2}$ 和 $f_{1-\alpha/2}$ 转变为等价的 f_α 和 $f_{1-\alpha}$ 值．第二个问题就有些困难，虽然也有完整的 F 表，以及 F 分布的计算程序，但我们想用表A.8来求解．因为这个限制，有两个可供参考的解法：**插值法**和**次下标 ν_1 和 ν_2 值法**．我们将在习题17.20中介绍插值法．这里我们用次下标值法．

表A.8中有 $\nu_2=14$，但没有 $\nu_1=14$．于是取次下标值，用 $\nu_1=12$ 和 $\nu_2=14$ 的分布，于是有

$$f_{\alpha/2,\nu_1,\nu_2}=f_{0.050/2,14,14}\approx f_{0.025,12,14}=3.05$$

为得到 $f_{\alpha/2,\nu_2,\nu_1}$ 的值，我们还是按习惯将 ν_2 假定为"ν_1"(见例17.9)，因此还是用分子值 $\nu_1=12$，于是

$$f_{\alpha/2,\nu_2,\nu_1}=f_{0.050/2,14,14}\approx f_{0.025,12,14}=3.05$$

故 σ_1^2/σ_2^2 的近似95%置信区间为

$$\left[\frac{3.33}{3.24}\left(\frac{1}{3.05}\right),\ \frac{3.33}{3.24}(3.05)\right]=(0.337,\ 3.135)$$

17.21 方差比(σ_1^2/σ_2^2)的假设检验:参数(σ_1^2, σ_1, μ_1 和 σ_2^2, σ_2, μ_2)未知的正态分布总体的独立样本

和其他参数的假设检验一样,每个方差比的假设检验问题都包含 H_0 和 H_1,也有双侧检验和单侧检验之分.在这些检验中,零假设叙述为,方差比 σ_1^2/σ_2^2 等于一个特定值 δ_0(见 17.5 节),所以,H_0 可写成

$$H_0: \sigma_1^2/\sigma_2^2 = \delta_0$$

虽然 δ_0 可以是任意一个值,但我们只考虑 $\delta_0 = 1$(即方差相等)的情形,它将产生下列可能的统计假设.

对于双侧检验,

$$H_0: \sigma_1^2/\sigma_2^2 = 1, \quad H_1: \sigma_1^2/\sigma_2^2 \neq 1$$

也可写成

$$H_0: \sigma_1^2 = \sigma_2^2, \quad H_1: \sigma_1^2 \neq \sigma_2^2$$

对于右侧检验,

$$H_0: \sigma_1^2/\sigma_2^2 = 1, \quad H_1: \sigma_1^2/\sigma_2^2 > 1$$

或

$$H_0: \sigma_1^2 = \sigma_2^2, \quad H_1: \sigma_1^2 > \sigma_2^2$$

对于左侧检验,

$$H_0: \sigma_1^2/\sigma_2^2 = 1, \quad H_1: \sigma_1^2/\sigma_2^2 < 1$$

或

$$H_0: \sigma_1^2 = \sigma_2^2, \quad H_1: \sigma_1^2 < \sigma_2^2$$

对于正态分布总体的独立抽样,可用 F 统计量[(17.34)式]检验零假设,在这些条件下,如果已知 $H_0: \sigma_1^2/\sigma_2^2 = 1$ 为真,则

$$F = \frac{S_1^2}{S_2^2} \tag{17.46}$$

这个统计量有 $\nu_1 = n_1 - 1$ 和 $\nu_2 = n_2 - 1$ 的 F 分布,并可取特定值

$$f = \frac{s_1^2}{s_2^2} \tag{17.47}$$

F 统计量可用于确定所需的条件概率

$$P(\text{与 1 至少相差 } s_1^2/s_2^2 \text{ 的点估计值} \mid H_0 \text{ 为真})$$

因为它满足检验统计量的两条标准(见 16.3 节):(1) 它考虑到点估计值(s_1^2/s_2^2)与零假设的比(σ_1^2/σ_2^2)相比较;(2)在 H_0 为真假定下,它有已知的概率分布(自由度为 ν_1 和 ν_2 的 F 分布).

当用 $F = S_1^2/S_2^2$ 作为检验统计量时,对一给定样本,特定值记作 $f^* = s_1^2/s_2^2$. 于是,用 F 作为检验统计量的假设检验的 P 值(见 16.6 节)是:得到 $\nu_1 = n_1 - 1$ 和 $\nu_2 = n_2 - 1$ 的 F 分布中的一个值的条件概率,这个值在 H_0 为真时至少和 f^* 一样大.对于右侧检验,P 值为

$$P = P(F \geqslant f^* \mid H_0 \text{ 为真}) \tag{17.48}$$

对于左侧检验，P 值为

$$P = P(F \leqslant f^* \mid H_0 \text{ 为真}) \tag{17.49}$$

对于双侧检验，不能简单的给出 P 值，因为 F 分布不对称，我们用 16.26 节中 χ^2 分布的解，给出步骤为

(1) 确定 f^* 是大于还是小于 F 分布的均值[(17.39)式]；

(2) 如果 $f^* > \mu$，P 值为

$$P = 2[P(F \geqslant f^* \mid H_0 \text{ 为真})] \tag{17.50}$$

(3) 如果 $f^* < \mu$，P 值为

$$P = 2[P(F \leqslant f^* \mid H_0 \text{ 为真})] \tag{17.51}$$

例 17.11　对于取自两个正态分布总体的独立样本($n_1 = n_2 = 9$)，在 $\alpha = 0.05$ 下，用 F 统计量，作 H_0：$\sigma_1^2/\sigma_2^2 = 1$ 和(a)H_1：$\sigma_1^2/\sigma_2^2 > 1$，(b) H_1：$\sigma_1^2/\sigma_2^2 < 1$，(c)$H_1$：$\sigma_1^2/\sigma_2^2 \neq 1$ 的检验，临界值叙述决策规则是什么？

解 ☞　(a) 对于右侧检验，其中 $\nu_1 = \nu_2 = 8$：拒绝 H_0，若 $f^* > (f_{\alpha/2,\nu_1,\nu_2} = f_{0.05,8,8} = 3.44)$.

(b) 对于左侧检验，其中 $\nu_1 = \nu_2 = 8$：拒绝 H_0，若

$$f^* < \left(f_{1-\alpha,\nu_1,\nu_2} = f_{1-0.05,8,8} = \frac{1}{f_{0.05,8,8}} = \frac{1}{3.44} = 0.29\right)$$

(c) 对于双侧检验，其中 $\nu_1 = \nu_2 = 8$：拒绝 H_0，若

$$f^* > (f_{\alpha/2,\nu_1,\nu_2} = f_{0.05/2,8,8} = f_{0.025,8,8} = 4.43)$$

或

$$f^* < \left(f_{1-\alpha/2,\nu_1,\nu_2} = f_{1-0.05/2,8,8} = f_{1-0.025,8,8} = \frac{1}{f_{0.025,8,8}} = \frac{1}{4.43} = 0.23\right)$$

例 17.12　对例 17.11 中的独立样本，当 $n_1 = n_2 = 9$，样本方差为 $s_1^2 = 4.23$，$s_2^2 = 1.56$. 在 $\alpha = 0.05$ 下，用临界值决策规则作 H_0：$\sigma_1^2/\sigma_2^2 = 1$ 的右侧检验.

解 ☞　(1) H_0：$\sigma_1^2/\sigma_2^2 = 1$，　H_1：$\sigma_1^2/\sigma_2^2 > 1$；

(2) $\alpha = 0.05$；

(3) 由于我们假定独立样本取自正态分布，所以用 F 统计量[(17.46)式]. 于是右侧检验决策规则为：拒绝 H_0，若 $f^* > (f_{\alpha,\nu_1,\nu_2} = f_{0.05,8,8} = 3.44)$.

(4) F 统计量的值为

$$f^* = \frac{s_1^2}{s_2^2} = \frac{4.23}{1.56} = 2.7115 \text{ 或 } 2.71$$

由表 A.8，得 $\nu_1 = 8$，$\nu_2 = 8$ 的 P 值在 $0.05 < P < 0.10$ 范围中.

(5) 因为 $f^* < 3.44$，所以接受 H_0. 由 P 值在 $0.05 < P < 0.10$ 范围中可确认此结论.

17.22　何时检验方差齐性

对于独立样本，用 t 方法时，对两样本估计和假设检验假定了方差的齐性(见 17.7 和 17.8 节). 我们在 17.21 节看到，如何用**方差不同的 F 检验**来检验这个假定. 那么在每次用 t 方法前都要做此检验吗？

仅在可以确定独立样本是取自正态分布总体时，方差不同的 F 检验或**方差比检验**或简单说成 **F 检验**才能用于检验方差齐性. 这是因为 F 检验对正态性的偏离比 t 方法对方差的非齐性更敏感(即更不稳健). 然而如果至少可假定总体是单峰的，则在样本容量相等的条件下，F 检验因方差不同的影响达到极小.

如果满足正态总体的假定，并用 F 检验拒绝 H_0，则不能假定方差齐性. 如 17.14 节所述一样，下一步是试图进行变换. 如果变换后，总体是正态分布，并且用一个新的 F 检验接受 H_0，则可假定方差齐性，并用变换后的数据用 t 方法. 如果变换后，总体不再是正态的，或者仍是拒绝 H_0，则可用非参数方法(见第 20 章)来检验数据的其他方面.

当满足正态分布的假定，并用 F 检验接受 H_0 时，这并不能证明方差的齐性，但允许接受这个假定.

17.23　比率差(p_1-p_2)的最优估计量：独立大样本

假定现研究两个具有相同特征的无限大二项总体(见 13.21 节)——即两个总体的所有元素都可分为具有(成功元素)和不具有(失败因素)某种特征——的关系. 两总体中成功元素的比例分别记作 p_1 和 p_2 是未知的. 现想估计这两个未知比率差(p_1-p_2). 为此，在 Bernoulli 试验条件下，从每个总体中抽取独立随机样本——第一总体中为 n_1 个元素，其中有 y_1 个是成功元素；第二总体中为 n_2 个元素，其中有 y_2 个是成功元素.

我们知道，当分别考虑每个总体时，样本中的成功比率$\left(\bar{P}=\frac{Y}{n}\right)$作为一个随机变量是总体比率$p$ 的无偏估计量(见 15.13 节). 如果我们记$\left(\bar{P}_1=\frac{Y_1}{n_1}\right)$是 p_1 的无偏估计量，$\left(\bar{P}_2=\frac{Y_2}{n_2}\right)$是 p_2 的无偏估计量，则我们说(不加证明)p_1-p_2 的最优无偏估计量是

$$\bar{P}_1-\bar{P}_2=\frac{Y_1}{n_1}-\frac{Y_2}{n_2} \tag{17.52}$$

它可取特定值

$$\bar{p}_1-\bar{p}_2=\frac{y_1}{n_1}-\frac{y_2}{n_2} \tag{17.53}$$

对取自两个总体的任意给定的一对样本，特定值 $\bar{P}_1-\bar{P}_2=\bar{p}_1-\bar{p}_2$ 是 p_1-p_2 的一个点估计值.

17.24　比率差的理论抽样分布

由 13.23 节知，如果所有容量为 n 的样本是取自 Bernoulli 条件下的无限大二项总体，并且对每个样本计算成功的比率 $P=p$，则比率的理论抽样分布的均值 $\mu_{\bar{p}}=p$、方差 $\sigma_{\bar{p}}^2=\frac{pq}{n}$、标准差 $\sigma_{\bar{p}}=\sqrt{\frac{pq}{n}}$，其中 p 是总体中的成功比率，$q=1-p$. 而且，由 13.25 节知，如果 n 足够大$(np\geqslant 5$ 且 $nq\geqslant 5)$，理论抽样分布为近似正态分布.

现在我们略去证明说，对于 17.23 节中的抽样条件，如果在 Bernoulli 条件下，所有容量为 n_1 和 n_2 的可能独立样本组合是独立地取自两个二项总体，并且对每个组合计算成功比率的差 $\bar{p}_1-\bar{p}_2$，则**两个比率差的理论抽样分布**或**比率差的理论抽样分布**对于均值、方差、标准差有下列公式：

$$\mu_{\bar{p}_1-\bar{p}_2}=p_1-p_2 \tag{17.54}$$

$$\sigma_{\bar{p}_1-\bar{p}_2}^2=\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2} \tag{17.55}$$

$$\sigma_{\bar{p}_1-\bar{p}_2}=\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}} \tag{17.56}$$

其中，(17.56)式称为**比率差的标准误差**. 另外，当 n_1 和 n_2 都足够大时，这个理论抽样分布可

假定为近似正态分布.

那么,何时 n_1 和 n_2 为足够大呢？对于未知 p_1、p_2,一些统计著作上说,当 $n_1q_1>5$, $n_1\bar{p}_1>5$, $n_2\bar{p}_2>5$ 且 $n_2\bar{q}_2>5$ 时,n_1, n_2 为足够大.其他著作上认为,上面四个值应都大于10或20或30.还有最保守的著作上认为,当 n_1 和 n_2 都大于等于100时,才认为足够大(见15.16节).

还可以指出,在Bernoulli试验条件下取自两个二项总体大样本条件,随机变量

$$Z=\frac{(\bar{P}_1-\bar{P}_2)-(p_1-p_2)}{\sqrt{\dfrac{p_1q_1}{n_1}+\dfrac{p_2q_2}{n_2}}} \tag{17.57}$$

有近似于正态分布的分布.

17.25　二项总体比率差 (p_1-p_2) 的近似置信区间:独立大样本

用17.23和17.24节的结论,我们现在讨论两比率差的 $(1-\alpha)100\%$ 置信区间.由(12.18)式知,对于标准正态分布有

$$P(-z_{\alpha/2}\leqslant Z\leqslant z_{\alpha/2})=1-\alpha$$

将(17.57)式代入(12.18)式得

$$P\left[-z_{\alpha/2}\leqslant\frac{(\bar{P}_1-\bar{P}_2)-(p_1-p_2)}{\sqrt{\dfrac{p_1q_1}{n_1}+\dfrac{p_2q_2}{n_2}}}\leqslant z_{\alpha/2}\right]\approx 1-\alpha$$

不等式每项同乘以 $\sqrt{\dfrac{p_1q_1}{n_1}+\dfrac{p_2q_2}{n_2}}$, 减去 $(P_1-\bar{P}_2)$, 再同乘以 -1 得

$$P\left[(\bar{P}_1-\bar{P}_2)-z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}\leqslant p_1-p_2\leqslant(\bar{P}_1-\bar{P}_2)+z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}\right]\approx 1-\alpha \tag{17.58}$$

有了这个关系式,就可以直接给出近似的 $(1-\alpha)100\%$ 置信区间.然而,正如总体比率的单样本估计一样,计算近似置信区间需要知道两个被估计的总体比率 p_1 和 p_2.为解决这个问题,我们还是用与单样本估计同样的方法(见15.15节),用点估计解和保守解.

在点估计解中,如果 n_1 和 n_2 都"足够大"(见17.24节),则点估计值 $\bar{p}_1$, $\bar{q}_1$, $\bar{p}_2$ 和 $\bar{q}_2$ 都可以代替置信限中的 p_1, q_1, p_2 和 q_2.因此,这些置信限可写成四种标准近似区间的形式:

(1) $(\bar{p}_1-\bar{p}_2)-z_{\alpha/2}\sqrt{\dfrac{\bar{p}_1\bar{q}_1}{n_1}+\dfrac{\bar{p}_2\bar{q}_2}{n_2}}\leqslant p_1-p_2\leqslant(\bar{p}_1-\bar{p}_2)+z_{\alpha/2}\sqrt{\dfrac{\bar{p}_1\bar{q}_1}{n_1}+\dfrac{\bar{p}_2\bar{q}_2}{n_2}}$

(2) $L\approx(\bar{p}_1-\bar{p}_2)-z_{\alpha/2}\sqrt{\dfrac{\bar{p}_1\bar{q}_1}{n_1}+\dfrac{\bar{p}_2\bar{q}_2}{n_2}}$, $U\approx(\bar{p}_1-\bar{p}_2)+z_{\alpha/2}\sqrt{\dfrac{\bar{p}_1\bar{q}_1}{n_1}+\dfrac{\bar{p}_2\bar{q}_2}{n_2}}$

(3) $\left[(\bar{p}_1-\bar{p}_2)-z_{\alpha/2}\sqrt{\dfrac{\bar{p}_1\bar{q}_1}{n_1}+\dfrac{\bar{p}_2\bar{q}_2}{n_2}},\ (\bar{p}_1-\bar{p}_2)+z_{\alpha/2}\sqrt{\dfrac{\bar{p}_1\bar{q}_1}{n_1}+\dfrac{\bar{p}_2\bar{q}_2}{n_2}}\right]$

(4) $(\bar{p}_1-\bar{p}_2)\pm z_{\alpha/2}\sqrt{\dfrac{\bar{p}_1\bar{q}_1}{n_1}+\dfrac{\bar{p}_2\bar{q}_2}{n_2}}$

保守解应用了这个事实:即不顾 p 的值,可以证明 $p(1-p)=pq$ 的最大可能值是 $\dfrac{1}{4}$.因此,如果 n_1 和 n_2 不是"足够大",或者要求用最保守可能解,则可用 $\dfrac{1}{4}$ 代替置信限中的 p_1q_1 和 p_2q_2 项,这样上面的第四种形式变为

$$(\bar{p}_1-\bar{p}_2)\pm z_{\alpha/2}\sqrt{\frac{1}{4n_1}+\frac{1}{4n_2}} \tag{17.59}$$

对于无限总体抽样,或有放回的有限总体抽样,以及无放回有限总体抽样,当 $n\leqslant 0.05N$ 对两个总体都成立时,两种置信区间的形式都是正确的.如果在最后一种情形中,对两个总体都有 $n>0.05N$,则用有限总体修正因子(见15.17节),可以给出类似的置信限公式.

例 17.13 在 Bernoulli 试验条件下,从两个无限大二项总体抽取独立随机样本($n_1=200$, $n_2=200$),得到 $y_1=105$, $y_2=90$.用(a)点估计解,(b)保守解,给出 p_1-p_2 的近似95%置信区间.

解 (a) p_1、q_1,p_2,q_2 的点估计值是:

$$\bar{p}_1=\frac{y_1}{n}=\frac{105}{200}=0.525,\ \bar{q}_1=1-0.525=0.475$$

$$\bar{p}_2=\frac{y_2}{n_2}=\frac{90}{200}=0.450,\ q_2=1-0.450=0.550$$

因为 n_1 和 n_2 都满足样本容量的最保守要求($n\geqslant 100$),所以可以用近似正态方法.由表A.5,$z_{\alpha/2}=z_{0.05/2}=1.96$.于是,用点估计解得到的近似95%置信区间为

$$(\bar{p}_1-\bar{p}_2)\pm z_{\alpha/2}\sqrt{\frac{\bar{p}_1\bar{q}_1}{n_1}+\frac{\bar{p}_2\bar{q}_2}{n_2}}$$

$$(0.525-0.450)\pm 1.96\sqrt{\frac{(0.525\times 0.475)}{200}+\frac{(0.450\times 0.550)}{200}}$$

$$0.075\pm 0.0977$$

(b) 用保守解[(17.59)式]得

$$(0.525-0.450)\pm 1.96\sqrt{\frac{1}{4(200)}+\frac{1}{4(200)}}$$

$$0.075\pm 0.0980$$

17.26 两个二项总体比率差(p_1-p_2)的假设检验:独立大样本

每个比率差的假设检验问题都包括零假设 H_0 和对立假设 H_1,它们或是双侧的,或是单侧的.在这些检验中,零假设叙述了总体比率差 p_1-p_2 等于某特定值 δ_0(见17.5节).于是,H_0 可写成 $H_0:p_1-p_2=\delta_0$.尽管 δ_0 可取任意值,但这里仅考虑 $\delta_0=0$ 的情况,检验的可能统计假设有:对于双侧检验

$$H_0:p_1-p_2=0,\quad H_1:p_1-p_2\neq 0$$

也可写成

$$H_0:p_1=p_2,\quad H_1:p_1\neq p_2$$

对于右侧检验

$$H_0:p_1-p_2=0,\quad H_1:p_1-p_2>0$$

或

$$H_0:p_1=p_2,\quad H_1:p_1>p_2$$

对于左侧检验

$$H_0:p_1-p_2=0,\quad H_1:p_1-p_2<0$$

或

$$H_0:p_1=p_2,\quad H_1:p_1<p_2$$

现在我们需要找到一个检验零假设的统计量.由17.24节知,如果我们假定样本容量 n_1 和 n_2 足够大,则 Z 统计量[(17.57)式]将有近似标准正态分布的分布.于是可用它确定所需的条件概率

$$P(\text{与 0 至少相差 } \bar{p}_1 - \bar{p}_2 \text{ 的点估计值} \mid H_0 \text{ 为真})$$

由于它满足检验统计量的两个条件:(1)它考虑到比较点估计值($\bar{p}_1 - \bar{p}_2$)与零假设值($p_1 - p_2 = 0$);(2)在已知 H_0 为真假定下,它有已知的概率分布(标准正态).

为计算 Z 统计量的特定值 z^*,可用点估计解和保守解[见(17.25)节]来代替(17.57)式的分母.但当检验 $H_0: p_1 - p_2 = 0$ 时,建议用**合并估计解**.在这个解中假定了 H_0 为真,即 $p_1 = p_2 = p$ 且 p 的最优估计值是 $\bar{p}$,它为样本中的成功元素的总数除以样本中的元素总数:

$$\bar{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{n_1 \bar{p}_1 + n_2 \bar{p}_2}{n_1 + n_2} \tag{17.60}$$

由于 $q = 1 - p$,所以 q 的最优估计值是 $\bar{q} = 1 - \bar{p}$,设 $p_1 - p_2 = 0$,将这些估计值代入 Z 统计量,则得到两特定样本计算 Z 值的公式

$$z^* = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2 = 0)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}} = \frac{(\bar{p}_1 - \bar{p}_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{17.61}$$

所有适用于 $H_0: p_1 - p_2 = 0$ 的单样本检验的 P 值和临界值决策规则(见16.10节),同样适用于两样本检验.

例17.14　对例17.13中的条件,在 $\alpha = 0.05$ 下,用临界值决策规则作 $H_0: p_1 - p_2 = 0$ 的双侧检验.

解 ☞　(1) $H_0: p_1 - p_2 = 0$, $H_1: p_1 - p_2 \neq 0$;

(2) $\alpha = 0.05$;

(3) 因为独立样本足够大($n_1 = n_2 = 200$),所以我们用 Z 统计量,得到双侧检验决策规则为:拒绝 H_0,若 $z^* > (z_{\alpha/2} = z_{0.050/2} = 1.96)$ 或 $z^* < (-z_{\alpha/2} = -z_{0.050/2} = -1.96)$

(4) 由(17.60)式得

$$\bar{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{105 + 90}{200 + 200} = 0.4875$$

$$\bar{q} = 1 - \bar{p} = 1 - 0.4875 = 0.5125$$

由例17.13 可知

$$\bar{p}_1 = 0.525,\ \bar{p}_2 = 0.450$$

Z 统计量的值为

$$z^* = \frac{(0.525 - 0.450)}{\sqrt{(0.4875)(0.5125)\left(\frac{1}{200} + \frac{1}{200}\right)}} = 1.500,\text{ 或 } 1.50$$

由表A.5,近似 P 值[(16.4)式]为

$$P \approx 2[P(Z \geqslant (z^* = a) \mid H_0 \text{ 为真})] = 2[P(Z > 1.50)] = 2(0.5 - 0.4332) = 0.1336$$

(5) 因为 $-1.96 < z^* < 1.96$,所以接受 H_0.这个结果可由近似 P 值大于0.05的予以确认.

习题解答

均值差的置信区间和假设检验:标准差已知的正态分布总体的独立样本

17.1　对下列各例,首先说明样本是独立的还是成对的,然后给出解释:

(a) 现有两个地点,为决定在哪个地点建购物中心,开发商选取的两地家庭收入随机样本;

(b) 为检验一项训练计划是否能降低某诊所中一组患者的高血压,在训练前后分别测

量的每个病人的血压;

(c) 为确定雄鹿的前后腿是否一样长,在雄鹿的随机样本中测量(前后的)右腿的长度;

(d) 为确定雄鹿和雌鹿的前腿是否一样长,在雄鹿与雌鹿的随机样本中测量右前腿的长度;

(e) 为比较两种用于船体的防藤壶(一种附在船地上的甲壳动物)油漆,将两种油漆都用于一组船的船体的相邻部分,一年后计算油漆处所有的累积藤壶数.

解 ☞ (a) 独立样本;在一个地点选取的观测值不影响另一个地点观测值的概率.

(b) 成对样本;训练后的观测值不是随机选取的.

(c) 成对样本;一旦测量了一条腿长,则另一条腿的长度就不是随机确定的了.

(d) 独立样本;选取的雄鹿腿的长度不会影响选取雌鹿腿的长度的概率,反之亦然.

(e) 成对样本;一旦计算了一艘船上一种油漆处的藤壶数,则另一种油漆上的数目就不是随机确定的了.

17.2 一位森林学家想知道海拔高度对红杉树高度的影响.他测量了海平面上 $n_1=100$ 棵成树的高度(总体1,标准差已知为 $\sigma_1=30\text{ft}$),海拔3,000英尺(ft)的 $n_2=73$ 棵成树的高度(总体2,标准差已知为 $\sigma_2=45\text{ft}$).得到 $\bar{x}_1=320\text{ft}$, $\bar{x}_2=255\text{ft}$.问 $\mu_1-\mu_2$ 的95%置信区间是什么?

解 ☞ 由(17.3)式,置信区间为(见17.4节)

$$(\bar{x}_1-\bar{x}_2)\pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

代入样本值和 $z_{0.05/2}=1.96$,得

$$(320\text{ft}-255\text{ft})\pm 1.96\sqrt{\frac{(30\text{ft})^2}{100}+\frac{(45\text{ft})^2}{73}}$$

$$65\text{ft}\pm 1.96\times 6.0613\text{ft}$$

$$65\text{ft}\pm 11.9\text{ft}$$

17.3 独立随机样本取自均值 μ_1, μ_2 未知,标准差已知的两个正态分布总体.第一个总体,$\sigma_1=0.73$,样本值 $n_1=25$, $\bar{x}_1=7.3$.第二个总体,$\sigma_2=0.89$,样本值 $n_2=20$, $\bar{x}_2=6.7$.在 $\alpha=0.01$ 下,用 P 值决策规则作 $H_0:\mu_1-\mu_2=0$ 的右尾检验.

解 ☞ (1) $H_0:\mu_1-\mu_2=0$, $H_1:\mu_1-\mu_2>0$;

(2) $\alpha=0.01$;

(3) 用 Z 统计量[(17.6)式],$\alpha=0.01$,右侧决策规则为:拒绝 H_0,若 $P\leqslant 0.01$.

(4) Z 统计量的值为

$$z^*=\frac{(7.3-6.7)-0}{\sqrt{\frac{(0.73)^2}{25}+\frac{(0.89)^2}{20}}}=2.431,\text{ 或 }2.43$$

为确定 P 值,用(16.6)式

$$P=P[Z\geqslant(z^*=a)\mid H_0\text{ 为真}]$$

$$=(Z>2.43)=0.5-0.4925=0.0075$$

(5) 因为 $P<0.01$,所以拒绝 H_0,接受 H_1.于是可以说,在0.01的显著水平下,两样本均值差明显大于0,这也表明 μ_1 大于 μ_2.

均值差的置信区间和假设检验:标准差未知但相等的正态分布总体的独立小样本

17.4 为检测一种兴奋剂对老鼠迷宫训练的影响,一位心理学家检验了两组雄性老鼠.第一组为对照组,给每只老鼠一个简单的日迷宫训练,直到它们达到训练标准,即无错误的通过迷宫(没有错误返回),在每天训练后,给每只老鼠注射生理盐水.第二组是试验组,除

了在每天训练后给每只老鼠注射含有兴奋剂的生理盐水以外,其他都和第一组相同.所取得度量是在达到训练标准前,每只老鼠在训练中错误次数的总和.样本是独立的,每组中有 10 只老鼠,结果如下:第一组 $n_1=10$, $\bar{x}_1=19.5$, $s_1^2=3.59$; 第二组 $n_2=10$, $\bar{x}_2=15.8$, $s_2^2=3.37$. 假定两个条件下的总体都是正态分布,并且方差相等(见 17.6 节),计算 $\mu_1-\mu_2$ 的 98% 置信区间.

解 ☞　本题可利用 17.7 节中的 $(1-\alpha)100\%$ 置信区间来解. 首先,由(17.14)式计算估计的标准误差为

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{\left[\frac{(10-1)3.59+(10-1)3.37}{18}\right]\left(\frac{1}{10}+\frac{1}{10}\right)}=0.8343$$

已知 $(\bar{x}_1-\bar{x}_2)=19.5-15.8=3.7$. 由表 A.6, 对 $(\nu=n_1+n_2-2=10+10-2=18)$ 的 t 分布,有 $t_{\alpha/2,\nu}=t_{0.02/2,18}=2.552$. 因此,98% 的置信区间为

$$3.7\pm(2.552\times0.8343)$$

$$3.7\pm2.13$$

17.5 为检测某种激素对失眠的影响,一个在睡眠不规律诊所的医生给两组临睡前的病人服用不同剂量的激素,然后测量他们从服药到入睡(由脑电波确定)的时间. 第一组服用的是 5mg 的剂量,第二组服用的是 15mg 的剂量,样本是独立的. 结果为:第一组 $n_1=10$, $\bar{x}_1=14.8\text{min}$, $s_1^2=4.36\text{min}^2$;第二组 $n_2=12$, $\bar{x}_2=10.2\text{min}$, $s_2^2=4.66\text{min}^2$. 假定两个条件下的总体是正态分布,并且有同方差(见 17.6 节),计算 $\mu_1-\mu_2$ 的 99% 置信区间.

解 ☞　本题可用 17.7 节中的 $(1-\alpha)100\%$ 置信区间求解. 首先,由(17.14)式计算的估计的标准误差为

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{\left[\frac{(10-1)4.36\text{min}^2+(12-1)4.66\text{min}^2}{20}\right]\left(\frac{1}{10}+\frac{1}{10}\right)}=0.9108\text{min}$$

已知 $(\bar{x}_1-\bar{x}_2)=14.8\text{min}-10.2\text{min}=4.6\text{min}$. 由表 A.6, 对 t 分布,当 $\alpha=0.01$, $(\nu=n_1+n_2-2=10+12-2=20)$ 时,有 $t_{\alpha/2,\nu}=t_{0.01/2,20}=2.845$. 于是,99% 的置信区间为

$$4.6\text{min}\pm(2.845\times0.9108\text{min})$$

$$4.6\text{min}\pm2.59\text{min}$$

17.6 再次考虑例 17.3 和例 17.4 中的记忆力研究. 在检验前,心理学家已知下午 7 点看电影的一组(第一组)比早上 7 点看电影的一组(第二组)反映要好. 于是,在 $\alpha=0.01$ 下,用临界值决策规则作 $H_0:\mu_1-\mu_2=0$ 的右尾检验.

解 ☞　(1) $H_0:\mu_1-\mu_2=0$, $H_1:\mu_1-\mu_2>0$;

(2) $\alpha=0.01$;

(3) 由于可假定正态分布,方差齐性和样本独立,所以用 t 统计量[(17.17)式]. 由 $\alpha=0.01$、$\nu=28$ 及表 A.6 得,右尾检验决策规则为:拒绝 H_0,若 $t^*>(t_{0.01,28}=2.467)$.

(4) 因为 $t^*=2.418$(见例 17.4),P 值在 $0.01<P<0.025$ 范围内.

(5) 因为 $t^*<2.467$,所以接受 H_0. 这一结果可由 P 值在 $0.01<P<0.025$ 范围中予以确认. 于是和例 17.4 中 $\alpha=0.05$ 下的显著双侧结果不同,$\alpha=0.01$ 下的右尾检验,不能说明第一组比第二组效果好.

17.7 对习题 17.4 的迷宫训练研究,在 $\alpha=0.01$ 下,用临界值决策规则作 $H_0:\mu_1-\mu_2=0$ 的双侧检验.

解 ☞　(1) $H_0:\mu_1-\mu_2=0$, $H_1:\mu_1-\mu_2\neq0$;

(2) $\alpha=0.01$;

(3) 因为可以假定正态分布,方差齐性和样本独立,所以用 t 统计量[(17.17)式]. 由 $\nu=10+10-2=18$, $\alpha=0.01$ 及表 A.6 知,双侧检验的决策规则为:拒绝 H_0,若 $t^*>(t_{0.01/2,18}=2.878)$或 $t^*<$

$(-t_{0.01/2,18}=-2.878)$.

(4) 用习题 17.4 中的 $s_{\bar{x}_1-\bar{x}_2}=0.8343$，$t$ 统计量的值为

$$t^*=\frac{3.7-0}{0.8343}=4.4349 \text{或} 4.435$$

$\nu=18$ 时，P 值范围为 $P<(2\times0.0005)$.

(5) 因为 $t^*>2.878$，所以拒绝 H_0，接受 H_1. 这一结果可由 P 值在 $P<0.001$ 范围中予以确认. 因此在 0.01 的显著水平下，两样本差与零有明显的不同，而且结果说明，试验组(第二组)平均起来比第一组试验次数要少些.

17.8 在进行习题 17.7 中的检验前，已知为达到标准平均地所试验组(第二组)试验次数要少. 于是应该在 $\alpha=0.01$ 下，用临界值决策规则作 $H_0:\mu_1-\mu_2=0$ 的右尾检验.

解☞ (1) $H_0:\mu_1-\mu_2=0$，$H_1:\mu_1-\mu_2>0$；

(2) $\alpha=0.01$；

(3) 用 t 统计量[(17.17)式]，由 $\nu=18$，$\alpha=0.01$ 以及表 A.6 知，右侧检验的决策规则为：拒绝 H_0，若 $t^*>(t_{0.01,18}=2.552)$.

(4) 和习题 17.7 一样，$t^*=4.435$，但现在对于这个右侧检验，$\nu=18$ 时，P 值范围为 $P<0.0005$.

(5) 因为 $t^*>2.552$，所以拒绝 H_0，接受 H_1. 这个结果可由 P 值在 $P<0.0005$ 范围中予以确认. 因此在 0.01 的显著水平下，平均说来注入兴奋剂的试验组试验次数要少些.

17.9 对习题 17.5 中的激素研究，在 $\alpha=0.02$ 下，用临界值决策规则作 $H_0:\mu_1-\mu_2=0$ 的双侧检验.

解☞ (1) $H_0:\mu_1-\mu_2=0$，$H_1:\mu_1-\mu_2\neq0$；

(2) $\alpha=0.02$；

(3) 因为可以假定正态分布、方差齐性以及样本独立，所以用 t 统计量[(17.17)式]. 由($\nu=n_1+n_2-2=10+12-2=20$)，$\alpha=0.02$ 及表 A.6 知，双侧检验的决策规则为：拒绝 H_0，若 $t^*>(t_{0.02/2,20}=2.528)$或 $t^*<(-t_{0.02/2,20}=-2.528)$.

(4) 用习题 17.5 中的 $s_{\bar{x}_1-\bar{x}_2}=0.9108\text{min}$ 得，t 统计量的值为

$$t^*=\frac{4.6\text{min}-0\text{min}}{0.9108\text{min}}=5.0505 \text{，或} 5.051$$

$\nu=20$ 时，P 值范围为 $P<(2\times0.0005)$或 $P<0.001$.

(5) 因为 $t^*>2.528$，所以拒绝 H_0，接受 H_1. 这一结果可由 P 小于 0.001 予以确认. 因此在 0.02 的显著水平下，两个样本差与零有显著的不同，而且结果说明 15mg 组比 5mg 组入睡要快.

17.10 在进行习题 17.9 中的检验前，已知 15mg 组(第二组)平均入睡比 5mg 组要快，于是应该在 $\alpha=0.025$ 下，用临界值决策规则作 $H_0:\mu_1-\mu_2=0$ 的右侧检验.

解☞ (1) $H_0:\mu_1-\mu_2=0$，$H_1:\mu_1-\mu_2>0$；

(2) $\alpha=0.025$；

(3) 用 t 统计量[(17.17)式]，由 $\nu=20$，$\alpha=0.025$ 以及表 A.6 知，右侧检验的决策规则为：拒绝 H_0，若 $t^*>(t_{0.025,20}=2.086)$.

(4) 和习题 17.9 一样，$t^*=5.051$，但现在对于这个右侧检验，$\nu=20$ 时，P 值范围为 $P<0.0005$.

(5) 因为 $t^*>2.086$，所以拒绝 H_0，接受 H_1. 这一结果可由 P 值小于 0.0005 予以确认. 因此在 0.025 的显著水平下，我们可以说 15mg 组比 5mg 组入睡要快.

均值差的置信区间和假设检验：标准差未知的任意总体分布的独立大样本

17.11 为试图养殖一种大鲤鱼，生物学家已经发现如何用遗传工程(转基因)养殖鲤鱼，使它们具有大鱼种的生长激素基因. 她养了 61 条这样的转基因鲤鱼，等到一年后称它们的重量，得到结果：$\bar{x}_1=1.69\text{lb}$，$s_1=0.254\text{lb}$. 同时，与转基因鱼同样的条件下，她还养了 61 条普通鲤鱼，一年后得到结果：$\bar{x}_2=1.57\text{lb}$，$s_2=0.250\text{lb}$. 假定两个总体都是正

态分布(对于这样的称重结果,这是典型的),并且方差齐性.问 $\mu_1-\mu_2$ 的95%置信区间是什么?

解 ☞　在本题的条件下,由表A.6有精确 t 解,并有17.7节中的置信区间.首先用(17.14)式计算 $s_{\bar{x}_1-\bar{x}_2}$

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{\left[\frac{(61-1)(0.254\text{lb})^2+(61-1)(0.250\text{lb})^2}{61+61-2}\right]\left(\frac{1}{61}+\frac{1}{61}\right)}=0.0456\text{lb}$$

已知 $(\bar{x}_1-\bar{x}_2)=1.69\text{lb}-1.57\text{lb}=0.12\text{lb}$,由表A.6,对 $\nu=61+61-2=120$ 的 t 分布有:$t_{\alpha/2,\nu}=t_{0.05/2,120}=1.980$.于是,$\mu_1-\mu_2$ 的95%置信区间为

$$0.12\text{lb}\pm(1.980\times0.0456\text{lb})$$

$$0.12\text{lb}\pm0.090\text{lb}$$

17.12 一位机场管理人员让你估计一下,两条航线中哪一条更遵守他们的计划起飞时间.对每条航线你随机测量了30架飞机的计划起飞和实际起飞的时间差[分钟(min)].现不能假定时间总体是正态分布,或是方差齐性的.独立样本结果为:航线(1),$\bar{x}_1=12.4\text{min}$,$s_1=3.72\text{min}$;航线(2),$\bar{x}_2=11.7\text{min}$,$s_2=3.60\text{min}$.问 $\mu_1-\mu_2$ 的95%置信区间是什么?

解 ☞　虽然不能假定正态分布,但为大样本($n_1=30$ 且 $n_2=30$),于是,可以用(17.19)式得到近似的99%置信区间.首先,计算估计的标准误差(17.7式)

$$s_{\bar{x}_1-\bar{x}_2}=\sqrt{\frac{(3.72\text{min})^2}{30}+\frac{(3.60\text{min})^2}{30}}=0.945\text{min}$$

已知 $(\bar{x}_1-\bar{x}_2)=12.4\text{min}-11.7\text{min}=0.7\text{min}$.由表A.5:$z_{0.01/2}=2.575$.于是,$\mu_1-\mu_2$ 的近似99%置信区间为

$$0.7\text{min}\pm(2.575\times0.945\text{min})$$

$$0.7\text{min}\pm2.43\text{min}$$

17.13 在进行习题17.11的研究前,已知平均起来转基因鲤鱼一年后比普通鲤鱼要重.于是,对习题17.11中的条件,在 $\alpha=0.05$ 下,用临界值决策规则作 $H_0:\mu_1-\mu_2=0$ 的右侧检验.

解 ☞　(1) $H_0:\mu_1-\mu_2=0$,　$H_1:\mu_1-\mu_2>0$;

(2) $\alpha=0.05$;

(3) 在本题条件下可用 t 统计量[(17.17)式].由($\nu=n_1+n_2-2=61+61-2=120$),$\alpha=0.05$ 以及表A.6知,右侧检验的决策规则为:拒绝 H_0,若 $t^*>(t_{0.05,120}=1.658)$.

(4) t 统计量的值为

$$t^*=\frac{(1.69\text{lb}-1.57\text{lb})-0}{\sqrt{\left[\frac{(61-1)(0.254\text{lb})^2+(61-1)(0.250\text{lb})^2}{61+61-2}\right]\left(\frac{1}{61}+\frac{1}{61}\right)}}=2.6289,\text{或}2.630$$

由表A.6知,$\nu=120$ 时,P 值范围为 $0.0005<P<0.005$.

(5) 因为 $t^*>1.658$,所以拒绝 H_0,接受 H_1.这个结果可由 P 值在 $0.0005<P<0.005$ 范围中予以确认.因此在0.05的显著水平下,由明显的证据表明平均起来一年后转基因鱼要比普通鱼重.

17.14 对习题17.12中的条件,在 $\alpha=0.01$ 下,用临界值决策规则作 $H_0:\mu_1-\mu_2=0$ 的双侧检验.

解 ☞　(1) $H_0:\mu_1-\mu_2=0$,　$H_1:\mu_1-\mu_2\neq0$;

(2) $\alpha=0.01$;

(3) 在本题条件下,Z 统计量[(17.19)式]有近似标准正态分布.且是由 $n_1=30$,$n_2=30$,$\alpha=0.01$ 知,近似的双侧检验的决策规则为:拒绝 H_0,若 $z^*>(z_{0.01/2}=2.575)$ 或 $z^*<(-z_{0.01/2}=-2.575)$.

(4) Z 统计量的值为

$$z^* = \frac{(12.4\text{min} - 11.7\text{min}) - 0}{\sqrt{\frac{(3.72\text{min})^2}{30} + \frac{(3.60\text{min})^2}{30}}} = 0.741，或 0.74$$

由表 A.5 知，近似 P 值为［(16.4)式］

$$P \approx 2[P(Z \geqslant (z^* = a) \mid H_0 为真)] = 2[P(Z > 0.74)] = 2(0.5 - 0.2704) = 0.4592$$

(5) 因为 $-2.575 < z^* < 2.575$，所以接受 H_0. 这一结果可由 P 值大于 0.01 予以确认. 因此在 0.01 的显著水平下，没有明显证据表明两条航线在准时起飞方面有差别.

均值差的置信区间和假设检验:成对样本

17.15 致力于激素研究的医生(见习题 17.5 和 17.9)发现由于随机选择的偶然性，15mg 组的平均年龄比 5mg 组的平均年龄要轻. 担心这个外部变量(见上册 3.10 节)可能会影响结果(15mg 组入睡快)，于是决定用成对样本模型重新研究. 首先根据可能影响睡眠的年龄、性别、健康状况以及其他因素，选取了 12 对病人，然后将每对病人随机分派到 5mg 和 15mg 组. 和以前一样，对每个病人测量从服药到入睡的时间，变量 X_1 代表 5mg 组的时间，变量 X_2 代表 15mg 组的时间. 对每对值计算时间差 $d_i = x_{1i} - x_{2i}$(以分钟计)：$d_1 = 4.9$，$d_2 = 4.6$，$d_3 = 5.1$，$d_4 = 4.5$，$d_5 = 7.1$，$d_6 = 3.2$，$d_7 = 5.4$，$d_8 = 3.9$，$d_9 = 5.9$，$d_{10} = 4.6$，$d_{11} = 2.9$，$d_{12} = 4.7$. 由这些数据给出 μ_d(5mg 和 15mg 组时间差的均值)的 95% 置信区间. 假定差值的总体是正态分布.

解 ☞ 用 17.11 节的公式，第一步得到 $\sum d_i = 56.8\text{min}$, $\sum d_i^2 = 282.92\text{min}^2$. 下面，用这些值得到

$$\bar{d} = \frac{\sum d_i}{n} = \frac{56.8\text{min}}{12} = 4.73\text{min}$$

$$s_d = \sqrt{\frac{\sum d_i^2 - n\bar{d}^2}{n-1}} = \sqrt{\frac{282.92\text{min}^2 - [12 \times (4.73\text{min})^2]}{12-1}} = 1.1459\text{min}$$

$$s_{\bar{d}} = \frac{s_d}{n} = \frac{1.1459\text{min}}{\sqrt{12}} = 0.3308\text{min}$$

由本题条件，可用 17.11 节的精确 t 解：

$$\bar{d} \pm t_{\alpha/2,\nu} s_{\bar{d}}$$

其中，由表 A.6 得 $t_{0.05/2,11} = 2.201$. 因此，95% 的置信区间为

$$4.73\text{min} \pm (2.201 \times 0.3308\text{min})$$

$$4.73\text{min} \pm 0.728\text{min}$$

17.16 某制药公司中的一科学家发现，在老鼠的食物中加入蒲公英中的一种物质，初步研究认为可能减少老鼠体内的血浆胆固醇含量(以 mg/dl 记). 为进行这种影响的成对模型研究，她选取了 35 对雄鼠. 每一对都是 10 周大，并将每对中的一只随机分配到两组中：对照组(X_1)，喂养 20 天正常鼠食；试验组(X_2)，喂养 20 天的含有蒲公英物质的鼠食. 20 天后，测量所有老鼠血浆胆固醇含量，并对每对计算差(对照组减试验组)$d_i = x_{1i} - x_{2i}$. 结果为 $\bar{d} = -0.53\text{mg/dl}$，$s_d = 1.895\text{mg/dl}$. 求 μ_d(对照组和试验组胆固醇水平差的总体均值)的 95% 置信区间. 假定差值的总体是正态分布.

解 ☞ 因为$(n = 35) > 30$，所以用 17.12 节中的精确 t 解或近似 Z 解. 由于表 A.6 不能给出精确 t 值. 于是，用近似 Z 解：$\bar{d} \pm z_{\alpha/2} s_{\bar{d}}$. 其中 $\bar{d} = -0.53\text{mg/dl}$，$s_{\bar{d}} = \frac{s_d}{\sqrt{n}} = \frac{1.895\text{mg/dl}}{\sqrt{35}} = 0.3203\text{mg/dl}$.

由表 A.5 得：$z_{0.05/2} = 1.96$. 于是，近似的 95% 置信区间为

$$-0.53\text{mg/dl} \pm (1.96 \times 0.3203\text{mg/dl})$$

$$-0.53\text{mg/dl} \pm 0.628\text{mg/dl}$$

17.17 对习题 17.15 中的激素研究，在 $\alpha=0.05$ 下，用临界值决策规则作 $H_0:\mu_d=0$ 的右侧检验.

解☞ (1) $H_0:\mu_d=0$, $H_1:\mu_d>0$;

(2) $\alpha=0.05$;

(3) 由于样本是成对的，并且假定差值(d_i)总体是正态分布，所以用 t 统计量[(17.30)式]，由 $\alpha=0.05$, ($\nu=n-1=12-1=11$)以及表 A.6 知，右尾检验的决策规则为：拒绝 H_0，若 $t^*>(t_{0.05,11}=1.796)$.

(4) t 统计量的值为

$$t^*=\frac{4.73\text{min}-0}{0.3308\text{min}}=14.2987 \text{ 或 } 14.299$$

$\nu=11$ 时的 P 值范围为 $P<0.0005$.

(5) 因为 $t^*>1.796$，所以拒绝 H_0，接受 H_1. 这个结果可由 P 值小于 0.0005 予以确认. 因此在 0.05 的显著水平下，成对样本差与零间有显著不同，这一结果表明 15mg 组入睡比 5mg 组快. 这些结果与独立样本研究结果一致(见习题 17.9 和 17.10).

17.18 对习题 17.16 的胆固醇研究，在 $\alpha=0.05$ 下，用临界值决策规则作 $H_0:\mu_d=0$ 的双侧检验.

解☞ (1) $H_0:\mu_d=0$, $H_1:\mu_d\neq0$;

(2) $\alpha=0.05$;

(3) 因为样本是成对的，并且是大样本[$(n=35)>30$]，而且假定差值(d_i)总体是正态分布，所以用 17.12 节中的 t 或 Z 解. 同习题 17.16 中的置信区间一样，由于表 A.6 不能给出精确 t 值，于是用近似 Z 解. 在 $\alpha=0.05$ 时，由表 A.5 得到近似值知，双侧检验的决策规则为：拒绝 H_0，若 $z^*>(z_{0.05/2}=1.96)$ 或 $z^*<(-z_{0.05/2}=-1.96)$.

(4) Z 统计量的值为

$$z^*=\frac{-0.53\text{mg/dl}-0}{0.3203\text{mg/dl}}=-1.6537 \text{ 或 } -1.65$$

由表 A.5 得，近似 P 值为

$$P\approx2[P(Z\geqslant(z^*=a)\mid H_0\text{为真})]=2[P(Z>1.65)]=2(0.5-0.4505)=0.0990$$

(5) 因为 $-1.96<z^*<1.96$，所以接受 H_0. 这一结果可由 P 值大于 0.05 予以确认. 因此在 0.05 的显著水平下，没有明显的证据表明蒲公英物质影响总胆固醇水平.

F 分布的临界值

17.19 利用表 A.8 求 $f_{\alpha/2,\nu_1,\nu_2}=f_{0.010/2,4,6}$ 和 $f_{1-\alpha/2,\nu_1,\nu_2}=f_{1-0.010/2,4,6}$.

解☞ 首先，将 $f_{\alpha/2}$ 和 $f_{1-\alpha/2}$ 的值转换为等价的 f_α 的值. 于是有

$$f_{0.010/2,4,6}=f_{0.005,4,6}=12.03$$

用倒数法则[(17.43)式]和例 17.9 中的方法，得

$$f_{1-0.010/2,4,6}=f_{1-0.005,4,6}=\frac{1}{f_{0.005,6,4}}=\frac{1}{21.97}=0.0455 \text{ 或 } 0.05$$

17.20 用插值法(见例 17.10)求 $f_{\alpha,\nu_1,\nu_2}=f_{0.050,11,7}$.

解☞ $f_{0.050,11,7}$ 不在表 A.8 中，但我们可以得到 $f_{0.050,10,7}=3.64$, $f_{0.050,12,7}=3.57$. 由标准线性插值顺序(见例 14.15)，首先得到图 17-5 中的 ν_1 和 $f_{0.05,\nu_1,7}$ 两条平行线. 然后，由于 $\nu_1=11$ 在 ν_1 线的 $\nu_1=10$ 到 $\nu_1=12$ 的 $\frac{1}{2}$ 距离的位置(用叉表示)，因此 $f_{0.05,11,7}$ 也应该在 $f_{0.05,\nu_1,7}$ 线的 3.64

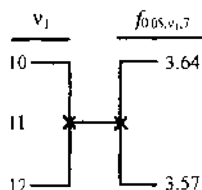


图 17-5

和 3.57 间的$\frac{1}{2}$处(也用义表示),于是近似的 F 值为

$$f_{0.050,11,7} \approx 3.64 - [0.5 \times (3.64 - 3.57 = 0.07)] \approx 3.605,\ \text{或}\ 3.60$$

17.21 利用表 A.8 求(a)$f_{\alpha,\nu_1,\nu_2} = f_{0.050,20,13}$、(b) $f_{1-\alpha,\nu_1,\nu_2} = f_{1-0.025,120,60}$,(c) $f_{1-\alpha,\nu_1,\nu_2} = f_{1-0.050/2,41,32}$. 如果临界值不在表 A.8 中,用次下标 ν_1 和 ν_2 方法求解(见例 17.10)

解 ☞ (a) $f_{0.050,20,13} = 2.46$.

(b) 由倒数法则[(17.43)式]有

$$f_{1-0.025,120,60} = \frac{1}{f_{0.025,60,120}} = \frac{1}{1.53} = 0.6536\ \text{或}\ 0.65$$

(c) 由于该值不在表 A.8 中,所以

$$f_{1-0.050/2,41,32} \approx f_{1-0.025,40,30} = \frac{1}{f_{0.025,30,40}} = \frac{1}{1.94} = 0.5155\ \text{或}\ 0.52$$

17.22 利用表 A.8 说明,当(a)ν_1 取 3 不变,而 ν_2 由 3 增加至 6 至 12;(b) ν_2 取 3 不变,而 ν_1 由 3 增加至 6 至 12;(c) ν_1,ν_2 同时由 3 增加至 6 至 12 时,$f_{0.050,\nu_1,\nu_2}$ 如何变化?

解 ☞ (a) $f_{0.050,3,3} = 9.28$, $f_{0.050,3,6} = 4.76$, $f_{0.050,3,12} = 3.49$

(b) $f_{0.050,3,3} = 9.28$, $f_{0.050,6,3} = 8.94$, $f_{0.050,12,3} = 8.74$

(c) $f_{0.050,3,3} = 9.28$, $f_{0.050,6,6} = 4.28$, $f_{0.050,12,12} = 2.69$

由这些结果可知,对相同的 α 水平,自由度越大,临界值越小.

17.23 利用表 A.8 说明:当 α 如下减小时:0.100,0.050,0.025,0.010,0.005,$f_{\alpha,10,10}$如何变化?

解 ☞

$$f_{0.100,10,10} = 2.32,\ f_{0.050,10,10} = 2.98,\ f_{0.025,10,10} = 3.72$$

$$f_{0.010,10,10} = 4.85,\ f_{0.005,10,10} = 5.85$$

由这些结果可知,对相同的 ν_1、ν_2,临界值随着 α 的减小而增大.

17.24 验证 F 分布与 t 分布、标准正态分布以及 χ^2 分布间的下列关系式,其中 $\alpha = 0.050$:

(a) $f_{\alpha,\nu_1=1,\nu_2} = (t_{\alpha/2,\nu=\nu_2})^2$, $\nu_2 = 5$; (b) $f_{\alpha,\nu_1=1,\nu_2=\infty} = (z_{\alpha/2})^2$; (c) $f_{\alpha,\nu_1,\nu_2=\infty} = \frac{\chi^2_{\alpha,\nu=\nu_1}}{\nu_1}$, $\nu_1 = 10$.

解 ☞ (a) 由表 A.8 得, $f_{0.050,1,5} = 6.61$

由表 A.6 得, $(t_{0.050/2,5})^2 = (2.571)^2 = 6.6100$, 或 6.61

(b) 由表 A.8 得, $f_{0.050,1,\infty} = 3.84$

由表 A.5 得, $(z_{0.050/2})^2 = (1.96)^2 = 3.8416$, 或 3.84

(c) 由表 A.8 得, $f_{0.050,10,\infty} = 1.83$

由表 A.7 得, $\frac{\chi^2_{0.050,10}}{10} = \frac{18.31}{10} = 1.831$, 或 1.83

方差比的置信区间和假设检验:参数未知的正态分布总体的独立样本

17.25 对习题 17.4 中的迷宫训练研究,计算 σ_1^2/σ_2^2 的 99% 置信区间.

解 ☞ 在习题 17.4 中,独立样本($n_1 = 10$, $s_1^2 = 3.59$, $n_2 = 10$, $s_2^2 = 3.37$)取自两个正态分布总体. 于是,用 17.20 节的$(1-\alpha)100\%$置信区间:

$$\left[\frac{s_1^2}{s_2^2}\left(\frac{1}{f_{\alpha/2,\nu_1,\nu_2}}\right),\ \frac{s_1^2}{s_2^2}(f_{\alpha/2,\nu_2,\nu_1})\right]$$

其中$\frac{s_1^2}{s_2^2}=\frac{3.59}{3.37}$. 由表 A.8 得 $f_{\alpha/2,\nu_1,\nu_2}=f_{0.010/2,9,9}=6.54$, $f_{\alpha/2,\nu_2,\nu_1}=f_{0.010/2,9,9}=6.54$. 因此 σ_1^2/σ_2^2 的 99%置信区间为

$$\left[\frac{3.59}{3.37}\left(\frac{1}{6.54}\right),\ \frac{3.59}{3.37}(6.54)\right]=(0.163,\ 6.967)$$

17.26　对习题 17.5 中的激素研究,计算 σ_1^2/σ_2^2 的 90%置信区间.

解　在习题 17.5 中,独立样本($n_1=10$, $s_1^2=4.36\text{min}^2$, $n_2=12$, $s_2^2=4.66\text{min}^2$)取自两个正态分布总体.于是,这里用 17.20 节中的$(1-\alpha)100\%$置信区间:其中$\frac{s_1^2}{s_2^2}=\frac{4.36\text{min}^2}{4.66\text{min}^2}$. 由表 A.8 得精确值:$f_{\alpha/2,\nu_1,\nu_2}=f_{0.100/2,9,11}=2.90$,以及用次下标 ν_1 和 ν_2 方法(见例 17.10)得到的近似值 $f_{\alpha/2,\nu_2,\nu_1}=f_{0.100/2,11,9}\approx f_{0.050,10,9}=3.14$. 于是 σ_1^2/σ_2^2 的近似 90%置信区间为

$$\left[\frac{4.36\text{min}^2}{4.66\text{min}^2}\left(\frac{1}{2.90}\right),\ \frac{4.36\text{min}^2}{4.66\text{min}^2(3.14)}\right]=(0.323,\ 2.938)$$

17.27　独立样本取自参数未知的两个正态分布总体,样本值为 $n_1=n_2=16$, $s_1^2=0.41$, $s_2^2=1.36$. 在 $\alpha=0.025$ 下,用临界值决策规则作 $H_0:\sigma_1^2/\sigma_2^2=1$ 的左侧检验.

解　(1) $H_0:\sigma_1^2/\sigma_2^2=1$,　$H_0:\sigma_1^2/\sigma_2^2<1$;

(2) $\alpha=0.025$;

(3) 由于可假定独立样本取自正态分布,所以用 F 统计量[(17.46)式],由表 A.8 得到左侧检验的决策规则为

$$拒绝\ H_0, 若\ f^* < \left(f_{1-0.025,15,15}=\frac{1}{f_{0.025,15,15}}=\frac{1}{2.86}=0.3497,\ 或\ 0.35\right)$$

(4) F 统计量的值为

$$f^*=\frac{s_1^2}{s_2^2}=\frac{0.41}{1.36}=0.3015,\ 或\ 0.30$$

P 值[(17.49)式]为 $P=P(F\leqslant 0.030|H_0\ 为真)$. 由(3)知,$f_{1-0.025,15,15}=0.35$. 由表 A.8 我们有

$$f_{1-0.01,15,15}=\frac{1}{f_{0.01,15,15}}=\frac{1}{3.52}=0.2841,\ 或\ 0.28$$

于是当 $f^*=0.30$ 时,P 值在 $0.01<P<0.025$ 范围内.

(5) 因为 $f^*<0.35$,所以拒绝 H_0,接受 H_1. 此结果可由 P 值在 $0.01<P<0.025$ 范围中予以确认.

17.28　在习题 17.4 的迷宫训练研究中,假定方差齐性(其中 $n_1=10, s_1^2=3.59$, $n_2=10$, $s_2^2=3.37$). 在 $\alpha=0.05$ 下,用临界值决策规则作这个假定的的双侧检验.

解　(1) $H_0:\sigma_1^2/\sigma_2^2=1$,　$H_0:\sigma_1^2/\sigma_2^2\neq 1$;

(2) $\alpha=0.05$;

(3) 由于假定独立样本取自正态分布,所以用 F 统计量[(17.46)式],由表 A.8 得双侧检验的决策规则为

$$拒绝\ H_0, 若\ f^* > (f_{0.05/2,9,9}=4.03)\ 或若$$

$$f^* < \left(f_{1-0.05/2,9,9}=\frac{1}{f_{0.05/2,9,9}}=\frac{1}{4.03}=0.2481,\ 或\ 0.25\right)$$

(4) F 统计量的值为

$$f^*=\frac{s_1^2}{s_2^2}=\frac{3.59}{3.37}=1.0653,\ 或\ 1.07$$

为确定 P 值范围,我们用 17.21 节中的方法. 首先得到均值

$$\mu=\frac{\nu_2}{\nu_2-2}=\frac{9}{9-2}=1.29$$

因为 $f^* < \mu$,所以 P 值[(17.51)式]为 $P=2[P(F\leqslant 1.07 | H_0$ 为真)].由表 A.8 得

$$f_{1-0.100,9,9}=\frac{1}{f_{0.100,9,9}}=\frac{1}{2.44}=0.4098$$

于是,P 值范围是 $P>(2\times 0.1)$或 $P>0.2$.

(5) 因为 $0.25<f^*<4.03$,所以接受 H_0.此结果可由 P 值大于 0.2 予以确认.因此在 0.05 显著水平下,没有明显证据表明方差不同,于是我们接受方差齐性的假定.

17.29 在习题 17.5 的激素研究中假定方差齐性(其中 $n_1=10$, $s_1^2=4.36\text{min}^2$, $n_2=12$, $s_2^2=4.66\text{min}^2$).在 $\alpha=0.01$ 下,用临界值决策规则作这个假定的双侧检验.

解 (1) $H_0:\sigma_1^2/\sigma_2^2=1$, $H_0:\sigma_1^2/\sigma_2^2\neq 1$;

(2) $\alpha=0.01$;

(3) 由于假定独立样本取自正态分布,所以用 F 统计量[(17.46)式].于是由 $\alpha=0.01$, $\nu_1=10-1=9$, $\nu_2=12-1=11$ 以及表 A.8,用次下标方法(见例 17.10)得到的近似双侧检验决策规则[见例 17.11(c)]为

$$\text{拒绝 } H_0\text{,若 } f^* > (f_{0.010/2,9,11}=5.54)$$

或若 $$f^* < \left(f_{1-0.010/2,9,11}=\frac{1}{f_{0.010/2,9,11}}\approx\frac{1}{f_{0.010/2,10,9}}=\frac{1}{6.42}=0.1558\text{, 或 }0.16\right)$$

(4) F 统计量的值为

$$f^*=\frac{s_1^2}{s_2^2}=\frac{4.36\text{min}^2}{4.66\text{min}^2}=0.9356\text{ 或 }0.94$$

用 17.21 节中的方法,求 P 值范围,我们有

$$\mu=\frac{\nu_2}{\nu_2-2}=\frac{11}{11-2}=1.2222\text{, 或 }1.22$$

因为 f^* 小于均值,所以 P 值[(17.51)式]为 $P=2[P(F\leqslant 0.94 | H_0$ 为真)].同样,由表 A.8 用次下标值方法得

$$f_{1-0.100,9,11}=\frac{1}{f_{0.100,9,11}}\approx\frac{1}{f_{0.100,10,9}}=\frac{1}{2.42}=0.4132\text{, 或 }0.41$$

于是可以说 P 值在 $P>(2\times 0.1)$或 $P>0.2$ 范围内.

(5) 因为 $0.16<f^*<5.54$,所以接受 H_0.此结果可由 P 值大于 0.2 予以确认.因此在 0.01 显著水平下,没有明显的证据表明方差不同,于是我们接受方差齐次的假定.

两个二项分布比率差的近似置信区间和假设检验:独立大样本

17.30 若想在一个大城市的两个地区中修建一个新的体育场,在每个区选取了 1,000 名登记的投票人进行民意测验,进而估计登记人中同意在所在区修建体育场的比率 p_1 和 p_2.结果为:第一区,572 人同意;第二区,533 人同意.由这些结果,用(a)点估计法和(b)保守解(见 17.25 节)给出 p_1-p_2 的 99%置信区间.

解 (a)由于 n_1 和 n_2 都满足最保守的样本容量要求($n\geqslant 100$),所以可以用近似正态方法(见 17.24 节).再者由于每个区中投票人都超过 100,000 人,$n\leqslant 0.05N$.因此,不需要有限总体修正因子.p_1, q_1, p_2, q_2 的点估计值为

$$\bar{p}_1=\frac{y_1}{n_1}=\frac{572}{1,000}=0.572,\ \bar{q}_1=1-\bar{p}_1=1-0.572=0.428$$

$$\bar{p}_2=\frac{y_2}{n_2}=\frac{553}{1,000}=0.553,\ \bar{q}_2=1-\bar{p}_2=1-0.553=0.447$$

由表 A.5 有 $z_{\alpha/2}=z_{0.01/2}=2.575$.用点估计法得到的近似 99%置信区间为

$$(\bar{p}_1-\bar{p}_2)\pm z_{\alpha/2}\sqrt{\frac{\bar{p}_1\bar{q}_1}{n_1}+\frac{\bar{p}_2\bar{q}_2}{n_2}}$$

$$(0.572-0.553)\pm 2.575\sqrt{\frac{(0.572\times 0.428)}{1,000}+\frac{(0.553\times 0.447)}{1,000}}$$

$$0.019\pm 0.0571$$

(b) 用保守解[(17.59)式]得

$$(0.572-0.553)\pm 2.575\sqrt{\frac{1}{4(1,000)}+\frac{1}{4(1,000)}}$$

$$0.019\pm 0.0576$$

17.31 现对1,000名高胆固醇患者进行一种降胆固醇新药的临床试验.六个月里,给其中的500人(第一组)每天服用一片含有标准剂量的药,对另外500人(第二组)每天服用一片**安慰药**(不含此药的相同药片).所选取的测量之一是研究期间的头疼现象,结果是:第一组6.2%的人感到头疼;第二组3.9%的人感到头疼.将这些百分数转换成比率,然后用 p_1 和 p_2 代表两个未知的总体比率.用(a)点估计解和(b)保守解给出 p_1-p_2 的95%置信区间.

解 (a) 由于 n_1 和 n_2 都满足最保守的样本容量需要($n\geq 100$),所以用近似正态方法(见17.24节).而且,由于试验组的假设总体很大,因此不需要有限总体修正因子. p_1, q_1, p_2, q_2 的点估计值为: $\bar{p}_1=0.062$, $\bar{q}_1=0.938$, $\bar{p}_2=0.039$, $\bar{q}_2=0.961$.于是,由表A.5有 $z_{\alpha/2}=z_{0.05/2}=1.96$.用点估计解得到的近似95%置信区间为

$$(\bar{p}_1-\bar{p}_2)\pm z_{\alpha/2}\sqrt{\frac{\bar{p}_1\bar{q}_1}{n_1}+\frac{\bar{p}_2\bar{q}_2}{n_2}}$$

$$(0.062-0.039)\pm 1.96\sqrt{\frac{(0.062\times 0.938)}{500}+\frac{(0.039\times 0.961)}{500}}$$

$$0.023\pm 0.0271$$

(b) 用保守解[(17.59)式]得

$$(0.062-0.039)\pm 1.96\sqrt{\frac{1}{4(500)}+\frac{1}{4(500)}}$$

$$0.023\pm 0.0620$$

17.32 对习题17.30中的体育场民意调查,在 $\alpha=0.01$ 下,用 P 值决策规则作 $H_0:p_1-p_2=0$ 的双侧检验.

解 (1) $H_0:p_1-p_2=0$, $H_1:p_1-p_2\neq 0$

(2) $\alpha=0.01$

(3) 由于独立样本足够大($n_1=n_2=1,000$),所以用 Z 统计量[(17.57)式].双侧检验的决策规则为:拒绝 H_0,若 $P\leq 0.01$.

(4) 由(17.60)式有

$$\bar{p}=\frac{y_1+y_2}{n_1+n_2}=\frac{572+553}{1,000+1,000}=0.5625$$

$$\bar{q}=1-\bar{p}=1-0.5625=0.4375$$

由习题17.30有

$$\bar{p}_1=0.572,\quad \bar{p}_2=0.553$$

Z 统计量[(17.61)式]的值为

$$z^*=\frac{(0.572-0.553)}{\sqrt{(0.5625)(0.4375)\left(\frac{1}{1,000}+\frac{1}{1,000}\right)}}=0.856,\text{ 或 }0.86$$

由表A.5知,近似 P 值为[(16.4)式]

$$P\approx 2[P(Z\geq(z^*=a)\mid H_0\text{为真})]=2[P(Z>0.86)]=2(0.5-0.3051)=0.3898$$

(5) 因为近似 P 值大于0.01,所以接受 H_0. 因此在0.01的显著水平下,没有证据表明两个地区的比率不同.

17.33 对习题17.31中的胆固醇研究,在 $\alpha=0.05$ 下,用临界值决策规则作 $H_0:p_1-p_2=0$ 的双侧检验.

解 ☞ (1) $H_0:p_1-p_2=0$, $H_1:p_1-p_2\neq 0$;

(2) $\alpha=0.05$;

(3) 由于独立样本足够大($n_1=n_2=500$),所以用 Z 统计量[(17.57)式],由表A.5知,双侧检验的决策规则为:拒绝 H_0,若 $z^*>(z_{0.05/2}=1.96)$或 $z^*<(-z_{0.05/2}=-1.96)$.

(4) 由(17.60)式有

$$\bar{p}=\frac{y_1+y_2}{n_1+n_2}=\frac{n_1\bar{p}_1+n_2\bar{p}_2}{n_1+n_2}=\frac{(500\times 0.062)+(500\times 0.039)}{500+500}=0.0505$$

$$q=1-\bar{p}=1-0.0505=0.9495$$

由习题17.31有

$$\bar{p}_1=0.062,\quad \bar{p}_2=0.039$$

Z 统计量[(17.61)式]的值为

$$z^*=\frac{(0.062-0.039)}{\sqrt{(0.0505)(0.9495)\left(\frac{1}{500}+\frac{1}{500}\right)}}=1.661,\ \text{或}\ 1.66$$

由表A.5知,近似 P 值为[(16.4)式]

$$P\approx 2[P(Z\geqslant(z^*=a)\mid H_0\text{为真})]=2[P(Z>1.66)]=2(0.5-0.4515)=0.0970$$

(5) 因为 $-1.96<P<1.96$,所以接受 H_0. 这一结果可由 P 值大于0.05予以确认. 因此在0.05的显著水平下,没有证据表明两组的比例不同.

17.34 在做习题17.31中的胆固醇试验前,已知第一组头痛比例比第二组的大. 因此应在 $\alpha=0.05$ 下,用临界值决策规则作 $H_0:p_1-p_2=0$ 的右尾检验.

解 ☞ (1) $H_0:p_1-p_2=0$, $H_1:p_1-p_2>0$;

(2) $\alpha=0.05$;

(3) 还是用 Z 统计量[(17.57)式],由表A.5得,右尾检验的决策规则为:拒绝 H_0,若 $z^*>(z_{0.050}=1.645)$.

(4) 由习题17.33有,$z^*=1.66$. 由表A.5知,近似 P 值为[(16.6)式]

$$P\approx[P(Z\geqslant z^*=a)\mid H_0\text{为真})]=P(Z>1.66)=(0.5-0.4515)=0.0485$$

(5) 因为 $z^*>1.645$,所以拒绝 H_0,接受 H_1. 这一结果可由近似 P 值小于0.05予以确认. 因此用单侧检验可以说在0.05的显著水平下,服药组头痛比率比安慰药组大.

补充习题

均值差的置信区间和假设检验:标准差已知的正态分布总体的独立样本

17.35 一生物学家研究两种蛇类总体,想知道它们形态上是否不同. 为此数了鳞片环数,并知两个总体是正态分布的,且有已知的标准差:第一组 $\sigma_1=1.4$,第二组 $\sigma_2=1.2$. 分别从两个总体中选取了 $n_1=35$, $n_2=40$ 的独立随机样本,得到 $\bar{x}_1=17.0$, $\bar{x}_2=16.3$. 问 $\mu_1-\mu_2$ 的95%置信区间是什么?

答案:0.7 ± 0.59

17.36 对习题17.35中的两种蛇类总体,在 $\alpha=0.05$ 下,作 $H_0:\mu_1-\mu_2=0$ 的双侧检验,说明接受还是拒绝 H_0,给出 H_0 为真下得到这个 z^* 值的概率.

答案:拒绝 H_0, 因为 $z^*=2.31>z_{\alpha/2}=1.96$; $P=0.021$

17.37 一市场调查员想知道市区家庭与乡村家庭每月购物次数是否不同. 假定两个总体中的购物频数服从正态分布,调查员选取了市区家庭($n_1=50$个家庭,$\sigma_1=2.3$次/月)和乡村家庭($n_2=50$个家庭,$\sigma_1=$

2.8 次/月)的独立随机样本,并得到市区家庭 $\bar{x}_1=8.6$ 次/月;乡村家庭 $\bar{x}_2=7.4$ 次/月.问 $\mu_1-\mu_2$ 的 95%置信区间是什么?

答案:1.2±1.00

17.38 对习题 17.37 中描述的市场调查,在 $\alpha=0.05$ 下,作 $H_0:\mu_1-\mu_2=0$ 的右尾检验,说明接受还是拒绝 H_0,给出 H_0 为真下得到这个 z^* 值的概率

答案: 拒绝 H_0,因为 $z^*=2.34>z_\alpha=1.645$; $P=0.0096$

均值差的置信区间和假设检验:标准差未知但假定相等的正态分布总体的独立小样本

17.39 一自然资源保护者想知道野外北极熊的出生率是否比动物园中北极熊的出生率低.他计算了 15 到 30 岁生育高峰期动物园总体($n_1=24$, $\bar{x}_1=19.1$, $s_1=2.3$)和野外阿拉斯加总体($n_2=18$, $\bar{x}_2=16.3$, $s_2=4.1$)的每头雌熊生产的幼熊数.假定这两个总体的出生率是正态分布,并且有相同的方差,问均值差的 95%置信区间是什么?

答案:2.8±2.01

17.40 在习题 17.39 的研究前,研究人员已经有理由相信野外北极熊的出生率低于动物园中的出生率,于是在 $\alpha=0.05$ 下,作 $H_0:\mu_1-\mu_2=0$ 的右尾检验.说明接受还是拒绝 H_0,并给出 H_0 为真下得到这个 t^* 值的近似概率为多少?

答案:拒绝 H_0,因为 $t^*=2.814>t_{0.05,40}=1.684$; $0.005<P<0.05$

17.41 一老年医学家想知道一天的各时间是否对老年人的学习能力有影响.他选取了 75 岁以上的老人的独立随机样本,并将他们分为两组.两组中的每个人都阅读心理学著作中的一章,然后做了关于此文的 50 道多项选择题.第一组($n_1=12$)在上午 8 点阅读并进行测验,第二组($n_2=12$)在下午 6 点阅读并进行测验.阅读能力是由试卷上的正确回答数来测量的.结果为:$\bar{x}_1=42$ 个正确,$s_1=2.0$,$\bar{x}_2=39$ 个正确,$s_2=3.3$.假定两组的学习度量是方差齐性的正态分布,问 $\mu_1-\mu_2$ 的 95%置信区间是什么?

答案:3.0±2.31

17.42 对习题 17.41 中的记忆力研究,在 $\alpha=0.05$ 下,作 $H_0:\mu_1-\mu_2=0$ 的双侧检验.说明接受还是拒绝 H_0,给出 H_0 为真下得到这个 t^* 值的近似概率为多少?

答案:拒绝 H_0,因为 $t^*=2.693>t_{0.05/2,22}=2.074$; $0.01<P<0.02$

均值差的置信区间和假设检验:标准差未知的任何分布的独立大样本

17.43 餐馆老板计划在另一社区开家分店,想知道在那里能有多少顾客,并与他的第一家餐馆的人数作一比较.他认为每月到餐馆就餐的人次数不服从正态分布,选取了两个社区顾客的独立随机样本,询问他们每月到餐馆就餐的次数.结果为:$n_1=120$, $\bar{x}_1=6.5$ 次/月,$s_1=2.1$ 次/月;$n_2=120$,$\bar{x}_2=5.8$ 次/月,$s_2=1.8$ 次/月.问 $\mu_1-\mu_2$ 的 95%置信区间是什么?

答案:0.7±0.49

17.44 对习题 17.43 中的餐馆顾客研究,在 $\alpha=0.05$ 下,作 $H_0:\mu_1-\mu_2=0$ 的双侧检验.说明接受还是拒绝 H_0,给出 H_0 为真下得到这个 z^* 值的近似概率为多少?

答案:拒绝 H_0,因为 $z^*=2.77>z_{0.05/2}=1.96$; $P\approx0.006$

17.45 一遗传学家想知道两个村子的雌臭虫总体的遗传情况是否不同.一个可能不同的度量是翅膀上的黑点数目,不服从正态分布.他从两个总体各取了 60 只臭虫的独立随机样本,得到 $\bar{x}_1=3.8$ 点,$s_1=1.2$ 点;$\bar{x}_2=3.6$ 点,$s_2=1.3$ 点.问 $\mu_1-\mu_2$ 的 95%置信区间是什么?

答案:0.2±0.45

17.46 对习题 17.45 中的雌臭虫研究,在 $\alpha=0.05$ 下,作 $H_0:\mu_1-\mu_2=0$ 的双侧检验.说明接受还是拒绝 H_0,给出 H_0 为真下得到这个 z^* 值的近似概率为多少?

答案:接受 H_0,因为 $-1.96<z^*=0.88<1.96$; $P\approx0.38$

均值差的置信区间和假设检验:成对样本

17.47 奶牛场主想知道一种新型的激素是否可以增加牛奶的产量.他取了 22 头奶牛的随机样本,并在服用激素前 2 周内每天记录其产奶量,在用激素喂养 6 个月后的 2 周内记录产奶量.对每只奶牛他用喂激素后的平均牛奶总量减去喂激素前的平均牛奶总量,得到:$\bar{d}=1.23$ 加仑,$s_d=0.68$ 加仑.假定差值总体是正态分布,求 μ_d 的 95%置信区间是什么?

答案:1.23gal±0.302gal

17.48 对习题 17.47 中的牛奶产量研究,在 $\alpha=0.01$ 下,作 $H_0:\mu_d=0$ 的右尾检验.说明接受还是拒绝 H_0,

给出 H_0 为真下得到这个 t^* 值的近似概率为多少?

答案:拒绝 H_0,因为 $t^*=8.484>t_{0.01,21}=2.518$; $P<0.005$

17.49 旅游鞋生产商想知道新型鞋(总体 1)是否比现在出售的鞋(总体 2)用得更久.他选取了 35 人的随机样本,并让他们一只脚上穿新型鞋(选哪只脚是随机的),另一只脚穿现型鞋.然后记录每只鞋达到一定坏损程度所走的路程.他计算了每人两只鞋的英里差(新型的减现型的),得到平均差值 $\bar{d}=5$ 英里,$s_d=5$ 英里.假定差值的总体是正态分布,用近似 Z 解给出 μ_d 的 95% 置信区间.

答案:5 英里 ± 1.7 英里

17.50 对习题 17.49 中的旅游鞋研究,在 $\alpha=0.05$ 下,做 $H_0:\mu_d=0$ 的右侧检验.说明接受还是拒绝 H_0,给出 H_0 为真下得 z^* 值的近似概率为多少?

答案:拒绝 H_0,因为 $z^*=5.916>z_{0.05}=1.645$; $P\approx 0$

F 分布的临界值

17.51 利用表 A.8 求 $f_{\alpha,\nu_1,\nu_2}=f_{0.010,24,33}$的近似值.用两种近似方法求解:插值法和次下标 ν_1, ν_2 法(见例 17.10 和习题 17.20).

答案:用插值法得:$f_{0.010,24,33}\approx 2.47-[0.3\times(2.47-2.29=0.18)]=2.4160$, 或 2.42.用次下标法得:$f_{0.010,24,33}\approx f_{0.010,24,30}=2.47$

17.52 利用表 A.8 求 $f_{\alpha/2,\nu_1,\nu_2}=f_{0.100/2,15,17}$.

答案:$f_{0.100/2,15,17}=f_{0.050,15,17}=2.31$

17.53 利用表 A.8 求 $f_{1-\alpha/2,\nu_1,\nu_2}=f_{1-0.010/2,12,24}$.

答案:$f_{1-0.010/2,12,24}=f_{1-0.005,12,24}=\dfrac{1}{f_{0.005,24,12}}=\dfrac{1}{4.43}=0.2257$ 或 0.23

17.54 利用表 A.8,由次下标 ν_1, ν_2 法(见例 17.10)求 $f_{\alpha,\nu_1,\nu_2}=f_{0.100,40,37}$.

答案:$f_{0.100,40,37}\approx f_{0.100,40,30}=1.57$

方差比的置信区间和假设检验:参数未知的正态分布的独立样本

17.55 一生物学家想知道海拔高度是否对某种蜻蜓体形有影响.为得到答案,首先他选取了低海拔地区的 $n_1=41$ 和高海拔地区的 $n_2=41$ 的独立随机样本,得到体长 $s_1^2=0.15\text{mm}^2$, $s_2^2=0.12\text{mm}^2$.假定两个总体的体长服从正态分布,问总体方差比 σ_1^2/σ_2^2 的 95% 置信区间是什么?

答案:(0.665,2.350)

17.56 对习题 17.55 中的蜻蜓研究,在 $\alpha=0.05$ 下,做 $H_0:\sigma_1^2/\sigma_2^2=1$ 的双侧检验.说明接受还是拒绝 H_0,给出 H_0 为真下得到这个 f^* 值的近似概率为多少?

答案:接受 H_0,因为 $f^*=1.25<f_{0.05/2,40,40}=1.88$; $P>0.20$

二项总体比率差的近似置信区间和假设检验:独立大样本

17.57 心理学家想知道自然科学家中的左撇子是否比人文科学家中的多.他在自然科学部门(总体 1)和人文部门(总体 2)各取了 160 名教授的随机样本.如果 p 为左撇子的比率,q 为右撇子的比率,结果为:$\bar{p}_1=0.32$, $\bar{q}_1=0.68$, $\bar{p}_2=0.28$, $\bar{q}_2=0.72$.用点估计法给出比率差 p_1-p_2 的近似 95% 置信区间.

答案:0.04 ± 0.100

17.58 对习题 17.57 中的左撇子研究,用保守解给出比率差的近似 95% 置信区间.

答案:0.04 ± 0.110

17.59 对习题 17.57 中的左撇子研究,在 $\alpha=0.05$ 下,做 $H_0:p_1-p_2=0$ 的双侧检验.说明接受还是拒绝 H_0,给出 H_0 为真下得到这个 z^* 值的近似概率为多少?

答案:接受 H_0,因为 $-z_{0.05/2}=-1.96<z^*=0.781<z_{0.05/2}=1.96$, $P\approx 0.435$

第十八章 多个样本的参数估计与假设检验

18.1 多个样本推断

从第十四章到第十七章,我们考察了单样本及双样本的参数估计与假设检验问题. 在本章,我们将进而介绍推断统计学的另一个非常重要而复杂的领域:**多个样本的估计与假设检验**. 我们在本章将专注于总体均值与总体方差的多个样本推断;在后面的第二十章,作为非参数方法讨论的一部分,我们还会处理有关总体比率的多个样本推断问题(见20.5节).

为明确多个样本推断与此前的推断问题间的差别,我们考察下面的例子,本章后面将多次提到此例:

我们正着手改良一类新品小麦,它将被送上太空,在由计算机控制的培养室中生长. 为了最充分地利用太空培养室,要求该小麦品种须具备两大优点:植株矮,产量高. 在经过了数年的杂交试验之后,假设我们最终得到了三个改良品种作为选择对象:品种1,品种2,与品种3 在接下来的针对这三个品种的试验中,于每个品种各取5株在培养室中加以培育,然后在其成熟期(发芽后的130天)测量各株小麦的高度(单位:英寸).

在前面各章所涉及的推断问题里,我们一次至多处理两个总体,而在多个样本场合中,我们将同时处理三个或更多的总体,在本例中即是如此:我们同时分析三个相互独立的总体,其中每个总体对应于小麦的一个品种,各总体分别具有均值 μ_1,μ_2,μ_3 以及方差 σ_1^2,σ_2^2 和 σ_3^2. 根据来自这些总体的样本,我们需要解决诸总体参数的估计与检验问题. 为此,前面已有的处理方法将不再适用,代之而来的是如下介绍的新方法:**方差分析**.

18.2 方差分析

方差分析(常记作ANOVA,或ANOV及AOV)是利用样本信息来检验关于 $k(k>2)$ 个总体的均值的假设的方法. 此方法由R.A.Fisher建立(见17.17节),如今它已成为推断统计学的一个广泛的领域. 在本章我们仅扼要叙述方差分析最基本的形式.

虽然此方法是针对关于总体均值的假设检验而建立的,可它的某些成分却可用于确定诸总体均值的置信区间(见18.16节及18.24节)及估计各个总体方差的差距(见18.17节及18.24节).

对于方差分析的分析对象——我们研究、测量的结果来说,一定有一个或多个客观存在的独立变量,它们可能会影响到我们要观测的. 不能独立变化的因变量,我们称之为"**因子**". 比如,在18.1节的例子中,有一个因子(小麦的品种)可能会影响到因变量(植株高度),因此在试验中,所讨论的因子的类别叫"**处理**"或"**水平**". 在此例中,因子分作三个水平:品种1、品种2及品种3. 因子可以处理任何有意义的测量水平(在本例中,因子是"名义水平"的)(见上册2.4节). 不过,在应用方差分析法时,所涉及的因变量的取值必须是间隔尺度或比例尺度(见上册2.6节及2.7节).

此方法通过探究样本数据的**变差来源**. 从而衡量各总体均值间的差异. 其本质为:判断样本均值的变差是由于"处理"的不同造成的(样本确实是来自均值不同的总体),还是纯粹由于对相同的总体取样的随机变异导致的.

18.3 单向、双向及多向方差分析

在上述小麦试验中,仅有一个独立变量或因子:小麦的品种. 因此可以说,此项研究的观测数据(植株高度)只按一个标准**分类**:小麦的品种. 因而,针对此组数据的方差分析称作**单向**

ANOVA 或**单因子 ANOVA**,**单类 ANOVA**.

但往往还有这种情况:为了从一项试验中获取更多的启示,该试验的设计要求同时考察两个或更多的相互独立的因子对于某因变量的影响. 若该设计与两个因子有关,则数据可以按两个标准来分类,因而称之为**双向 ANOVA**;若设计包含了三个因子,则对应着**三向 ANOVA**,依此类推. 一般地,若因子超过了两个,则相应的方差分析就称为**多向 ANOVA**.

对我们的小麦试验而言,可以有一个需要用到双向方差分析的设计:考察小麦品种与温度对植株高度的影响. 甚至还可以设计三向方差分析试验,只需在试验中再添加第三个因子即可(比如光照量或培养液中某种营养成份的量).

本章只是对方差分析做逻辑与步骤上的扼要介绍,故而仅介绍单向方差分析.

18.4 单向方差分析:固定效应、随机效应

一个单向方差分析被称作是**固定效应**的,如果其因子水平的选取是出于设计者的偏好,而不是随机选定的. 一个单向方差分析被称作是**随机效应**的,如果其因子的水平是从大量的因子处理中随机选取的. 若设计是固定效应的,则据此而得出的结论必须仅局限于所采用的那些特定的水平值;若设计是随机效应的,则其结论的适用范围会广泛些,它可以外推到所采用的水平的源头——因子处理上去的.

如上两种形式的方差分析在方法处理上是不同的,而在本章,我们仅研究固定效应的方差分析. 我们的小麦品种试验可以用于任何一种模型,然而,我们将假定所研究的三个小麦品种是我们唯一感兴趣的品种,因而采用固定效应 ANOVA.

18.5 单向固定效应方差分析:各种假定

对于一项试验(探索性的,受控的或自然进行的;见上册 3.9 节至 3.11 节),如果其结果可以用单向固定效应方差分析法加以分析,则该试验必须具备一个独立变量(因子)与一个因变量(见 18.3 节),并且因子水平的选取不是随机的,而是凭兴趣人为设定的. 此外,它还应满足如下条件:

(1) k 个独立样本分别取自 k 个间隔尺度或比例尺度的总体. 样本容量可以相同($n_1=n_2=\cdots=n_k=n$)或不同. (在 18.6 节至 18.13 节中,我们给出相等样本容量的方差分析,之后在 18.18 节至 18.22 节,给出更一般的方法,它对于样本容量相同或不同都适用).

(2) k 个总体都服从正态分布,分别具有均值 $\mu_1,\mu_2,\cdots,\mu_k$,及方差 $\sigma_1^2,\sigma_2^2,\cdots,\sigma_k^2$.

(3) k 个总体方差相同,即 $\sigma_1^2=\sigma_2^2=\cdots=\sigma_k^2$. 从而可假定存在 σ^2,它是各总体的共同方差.

18.6 样本容量相等时的单向固定效应方差分析:H_0 与 H_1

方差分析是检验关于总体均值假设的一种方法. 在单向固定效应方差分析中,总有一对统计假设:零假设(H_0)与备择假设(H_1),其中,零假设为:

$$H_0: \mu_1=\mu_2=\cdots=\mu_k$$

备择假设可以用不同方式表示:

H_1:k 个均值不全相同.

H_1:至少有两个均值不相等.

H_1:并非所有均值都相同.

H_1:H_0 不正确

一个正态分布可以由参数 μ 与 σ^2 完全确定(见 12.3 节). 从而,由前面关于正态分布及

公共方差 σ^2 的假定,我们真正所要检验的问题是:样本是否来自同一个总体.

18.7　样本容量相等时的单向固定效应方差分析:数据的整理

假设我们已经收集了一组数据. 在此项研究中,有一个具有 k 个水平的因子,和一个因变量. 我们进一步假设对于这个因变量通过观测获得了 k 个样本,每个样本对应着一个水平,样本容量均为 n,如此便得到了 nk 个观测数据. 我们以符号 x_{ij} 来表示观测值,其中指标 i 表示从 1 到 k 个水平,j 表示在第 i 个样本中的从 1 到 n 的 n 个特定的观测值. 在单向方差分析中,所有这 nk 个数据都被以典型的方式概括地整理在表 18.1 中. 表中的行表示样本,而列表示各数据在相应样本的观测值序列中的位置. 比如,表格的第一行展示了样本 1(对应水平 1)中的第 1 个至第 n 个观测值序列:从 x_{11} 到 x_{1n},表格最左边的一列数据表示的是这 k 个样本中每个样本的第 1 个观测值:从 x_{11} 到 x_{k1}.

表 18.1

因子	因变量					合计	平均
1	x_{11}	x_{12}	$\cdots$ x_{1j}	$\cdots$	x_{1n}	$T_{1.}$	$\bar{x}_{1.}$
2	x_{21}	x_{22}	$\cdots$ x_{2j}	$\cdots$	x_{2n}	$T_{2.}$	$\bar{x}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
i	x_{i1}	x_{i2}	$\cdots$ x_{ij}	$\cdots$	x_{in}	$T_{i.}$	$\bar{x}_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
k	x_{k1}	x_{k2}	$\cdots$ x_{kj}	$\cdots$	x_{kn}	$T_{k.}$	$\bar{x}_{k.}$
						$T_{..}$	$\bar{x}_{..}$

表格中各行数据的总和,在方差分析中称作"合计",并记作 T(T 需大写,尽管它不是变量),它们被记在各行右端的"合计"栏内. 比如,符号 $T_{1.}$,它带有下标(1.),表明它是第一个水平下的所有观测值之和,其计算公式为

$$T_{1.}=\sum_{j=1}^{n}x_{1j} \tag{18.1}$$

对第 i 个水平而言,$T_{i.}$ 的计算式为

$$T_{i.}=\sum_{j=1}^{n}x_{ij} \tag{18.2}$$

在合计栏的下方,符号 $T_{..}$ 表示的是所有的 nk 个样本观测值之和,其计算式为

$$T_{..}=\sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij} \tag{18.3}$$

式中双重求和的意义是:首先对每个水平下的所有 n 个样本作和,然后从 1 到 k 对所有水平作和.

表中各行数据的平均值列在最右边的"平均"栏内. 例如,记号 $\bar{x}_{1.}$ 表示的是水平 1 下的所有观测值的平均值,其计算式为

$$\bar{x}_{1.}=\frac{\sum_{j=1}^{n}x_{1j}}{n} \tag{18.4}$$

相应于第 i 个水平,其平均值 $\bar{x}_{i.}$ 的计算公式为

$$\bar{x}_{i.}=\frac{\sum_{j=1}^{n}x_{ij}}{n} \tag{18.5}$$

在平均栏的下方，记号 $\bar{x}_{..}$ 表示的是所有的 nk 个观测值的平均值，称作"**总平均**"(见上册 6.10 节)，其计算公式如下：

$$\bar{x}_{..}=\frac{\sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij}}{nk} \tag{18.6}$$

例 18.1 考虑我们提到过的那个例子，假定已经对小麦的每个品种各种植了 5 株，在成熟期时测量各株的高度(精确到英寸，1 英寸 =2.54 厘米)，下面是测量的结果：品种 1(17,18,17,16,18)；品种 2(21,19,18,19,20)；品种 3(18,21,20,21,19)．试将这些数据整理成表 18.1 那样的概括表格．

解 ☞ 所求表格如表 18.2，其中已给出合计值与平均值：

表 18.2

小麦品种	高度(英寸)	合计	平均
1	17 18 17 16 18	86 英寸	17.2 英寸
2	21 19 18 19 20	97 英寸	19.4 英寸
3	18 21 20 21 19	99 英寸	19.8 英寸
		282 英寸	18.8 英寸

18.8 样本容量相等时的单向固定效应方差分析：基本原理

对于我们所讨论的单向固定效应方差分析而言，已知所得到的容量彼此相同的样本来自 k 个独立的具有共同方差 σ^2 的正态总体；而且，还知道我们将以某种方式利用这些样本的信息来检验零假设 $H_0:\mu_1=\mu_2=\cdots=\mu_k$．

检验步骤如下：根据数据确定 σ^2 的两个独立的点估计，然后再利用 F 检验来比较它们(见 17.21 节)．第一个估计作为比值 F 的分母，称之为**样本内估计或误差估计**，并记作 S_W^2，其计算过程如下：分别对每个样本计算方差，然后将这些方差取平均．σ^2 的这种估计刻画了**样本内部的随机变差或随机误差**．

σ^2 的第二个估计作为比值 F 的分子，称之为**样本间估计**，记作 S_A^2，它可由如下方法确定：首先计算各样本平均值 $\bar{x}_{i.}$ 关于总平均值 $\bar{x}_{..}$ 的标准差，由于此标准差就是平均值的标准误差 $s_{\bar{x}}$，而 $s_{\bar{x}}=\frac{s_A}{\sqrt{n}}$(见 14.3 节)，故 $s_A=\sqrt{n}s_{\bar{x}}$，从而 $s_A^2=ns_{\bar{x}}^2$ 这个估计也刻画了样本内部的随机变差，而且它还反映了由于水平的不同而可能导致的样本均值间的变差．

若零假设为真，则由抽样数据而得到的 F 的一个具体值：

$$f^*=\frac{s_A^2}{s_W^2} \tag{18.7}$$

它应该接近于 1.0．另一方面，若零假设不真，则此数值应显著地大于 1.0．单向方差分析的基本原理是：将 σ^2 的一个对零假设成立与否很敏感的估计(s_A^2)与另一个对零假设不敏感的估计(s_W^2)进行比较．我们知道，即使是从同一个总体抽样，由于抽样的随机性，所得结果也会有所波动，方差分析检验法则将能回答如下的疑问：样本均值间的变差是纯粹由这种随机波动造成的？还是此变差实在太大以致于不能简单地归因于抽样的随机波动呢？

18.9 *SST = SSA + SSW*

若将表 18.1 中的所有 k 个样本看成是一个含有 nk 个观测值的**总样本**，则此总样本的变差可以由一个名为"**总平方和**"(见上册 7.5 节)的量来刻划，记之为 SST，其计算式为

$$SST=\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{..})^2 \tag{18.8}$$

它是所有的 nk 个值与总平均[(18.6)式]的差的平方和. 所有的方差分析都是基于这样的一个 SST,并且按数据变差的不同来源用数学方法将它分解为若干部分. 对于单向固定效应方差分析而言,可以用数学方法将 SST 作如下分解:

$$SST = SSA + SSW \tag{18.9}$$

即

$$\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{..})^2 = n\sum_{i=1}^{k}(\bar{x}_{i.}-\bar{x}_{..})^2 + \sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})^2 \tag{18.10}$$

式中的第一项

$$SSA = n\sum_{i=1}^{k}(\bar{x}_{i.}-\bar{x}_{..})^2 \tag{18.11}$$

称作**"组间平方和"**或**样本间平方和**,**处理间平方和**及**因子间平方和**. 它是 18.8 节的计算式中的分子 s_A^2. 我们曾指出,s_A^2 作为 σ^2 的样本间估计,它是样本内部随机变差以及因处理不同而可能导致的样本均值变差的一种度量.

式中的第二项

$$SSW = \sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})^2 \tag{18.12}$$

称为**"组内平方和"**(或**样本内平方和**,**误差平方和**). 它是方差的表达式中的分母 s_W^2(见上册 7.7 节). 我们也说过,s_W^2 作为公共方差 σ^2 的样本内估计,是各个样本内部围绕各自平均值的变差的一种度量,它是独立于处理条件的.

(18.9)式与(18.10)式称为**平方和分解式**(见上册 1.12 节),由此式可知,SSW 显然可以通过计算 SST 与 SSA 而间接地得到

$$SSW = SST - SSA \tag{18.13}$$

例 18.2　对表 18.2 中的数据,首先利用(18.8),(18.11)及(18.12)各式计算 SST,SSA 与 SSW;然后利用所得结果验证(18.9),(18.10)与(18.13)式的正确性.

解 ☞

$$SST = \sum_{i=1}^{3}\sum_{j=1}^{5}(x_{ij}-\bar{x}_{..})^2 = 34.4\ 英寸^2$$

$$SSA = 5\sum_{i=1}^{3}(\bar{x}_{i.}-\bar{x}_{..})^2 = 19.6\ 英寸^2$$

$$SSW = \sum_{i=1}^{3}\sum_{j=1}^{5}(x_{ij}-\bar{x}_{i.})^2 = 14.8\ 英寸^2$$

验证(18.9)式和(18.10)式

$$34.4\ 英寸^2 = 19.6\ 英寸^2 + 14.8\ 英寸^2$$

验证(18.13)式

$$14.8\ 英寸^2 = 34.4\ 英寸^2 - 19.6\ 英寸^2$$

18.10　*SST* 与 *SSA* 的计算公式

在上册的第 7 章里,介绍了方差及标准差的计算公式(见 7.6,7.8,7.9,7.10,7.12 及7.13 节),对于 18.9 节中的各平方和,也有着便于运用的等价的代数表达式,这些计算式利用了表 18.1 中的合计与平均. 相应于(18.8)式,有计算公式

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n}x_{ij}^2 - \frac{T_{..}^2}{kn} \tag{18.14}$$

对于(18.11)式,有计算公式

$$SSA = \frac{\sum_{i=1}^{k} T_i^2}{n} - \frac{T^2}{kn} \tag{18.15}$$

然后再利用(18.13)式,便可求出 SSW.

例 18.3 对于表 18.2 中的数据,我们在例 18.2 中求出了 SST,SSA 与 SSW. 试利用公式(18.13),(18.14)与(18.15)式将各值重新计算.

解

$$SST = \sum_{i=1}^{3}\sum_{j=1}^{5} x_{ij}^2 - \frac{T^2}{kn} = 5,336 \text{ 英寸}^2 - (79,524 \text{ 英寸}^2/15) = 34.4 \text{ 英寸}^2$$

$$SSA = \frac{\sum_{i=1}^{3} T_i^2}{n} - \frac{T^2}{kn} = (26,606 \text{ 英寸}^2/5) - (79,524 \text{ 英寸}^2/15) = 19.6 \text{ 英寸}^2$$

$$SSW = SST - SSA = 34.4 \text{ 英寸}^2 - 19.6 \text{ 英寸}^2 = 14.8 \text{ 英寸}^2$$

18.11 自由度与均方

在 18.8 节中我们指出,在方差分析里可以利用比值 F[(18.7)式]的具体观测值

$$f^* = \frac{s_A^2}{s_W^2}$$

来检验零假设 $H_0: \mu_1 = \mu_2 = \cdots = \mu_k$. 式中的分子是公共方差 σ^2 的样本间估计,分母是 σ^2 的样本内估计. 然后在 18.9 节中指出了平方和 SST 可以分解为分子 s_A^2(SSA)与分母 s_W^2(SSW)之和.

为了完成比值 F 的构造,现在需要确定这两个方差估计量的分母. 为此,我们回到总体方差及样本方差的初始定义. 上册的(7.12),(7.13)式定义总体方差为

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N} = \frac{SS}{N}$$

将平方和除以总体容量就得到了总体方差. 可是对于样本方差而言,为了得到总体方差的无偏估计,必须将样本平方和除以自由度(记为 df):$n-1$[见上册的(7.16)式及 14.17 节].

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{SS}{df}$$

由此类推,为得到总体方差的无偏估计,必须将样本平方和除以一个恰当的自由度数值. 在方差分析里,这种方差估计称为**"均方"**(**组平均离差平方和的平均**的简称).

在本书中,我们已经看到数理统计里的自由度 df(或 v)是概率分布的参数(见 14.15,15.2与 17.7 节). 而且我们还指出,自由度可以解释为计算一个统计量时所用到的可自由取值的变元的个数(见 14.17 节).

利用这第二种解释,我们看到,对于样本间估计量,其自由度为 $k-1$. 这是因为在 SSA 中共含有 k 个离差的平方,而此前已算出了 $\bar{x}_{..}$,从而为进一步得到 SSA 的值,只剩下 $k-1$ 个离差平方是不受任何约束的. 将 SSA[(18.11)式]除以 $k-1$,就得到了总体方差的**样本间均方估计**,记为 MSA

$$MSA = \frac{SSA}{k-1} = \frac{n\sum_{i=1}^{k}(\bar{x}_{i.} - \bar{x}_{..})^2}{k-1} \tag{18.16}$$

由此可知,正如 18.8 节中已指出的那样,MSA 实际上就是 $ns_{\bar{x}}^2$. MSA 作为 σ^2 的估计,当 H_0

成立时，它是无偏的，而当 H_1 成立时，它是偏大的.

对于样本内估计量，相应自由度为 $k(n-1)$ 或 $kn-k$，原因在于：针对每个样本要先计算其平均值，因而为得到该样本的平方和，就只有 $n-1$ 个关于平均值的离差平方是自由的. 又由于样本共有 k 个，则与 SSW[(18.12)式]相应的自由度即为 $k(n-1)$. 将 SSW 除以 $k(n-1)$，就得到了 σ^2 的**样本内均方估计**，记之为 MSW

$$MSW=\frac{SSW}{k(n-1)}=\frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})^2}{k(n-1)} \tag{18.17}$$

不难看出，对每个样本 i，有

$$\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i.})^2=(n-1)s_i^2$$

因此，我们在 18.8 节曾提到的事实是成立的，即 MSW 实际上就是对 k 个样本方差(见 17.6 节)所作的平均. 不论 H_0 成立与否，MSW 都是 σ^2 的无偏估计.

最后，我们可以由总平方和 SST[(18.8)式]构造出 σ^2 的第三个估计，只要将 SST 除以其自由度 $kn-1$ 即可. 这个自由度是正确的，因为我们是在算出了 $\bar{x}$ 后再计算 SST 的，从而余下的可自由变化的离差平方为 $kn-1$ 个. 将 SST 除以 $kn-1$ 就得到了总体方差的**总均方估计**，记之为 MST

$$MST=\frac{SST}{kn-1}=\frac{\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{..})^2}{kn-1} \tag{18.18}$$

应该指出，相应于平方和分解式[(18.9)与(18.10)式]，各平方和的自由度间也存在相应的关系式

$$df(\text{总的平方})=df(\text{组间平方和})+df(\text{组内平方和}) \tag{18.19}$$

$$kn-1=(k-1)+k(n-1) \tag{18.20}$$

$$=k-1+kn-k=kn-1$$

例 18.4　利用例 18.2 与例 18.3 中的 SSA、SSW 及 SST 的值，计算 MSA，MSW 及 MST.

解 ☞　由(18.16)式，有

$$MSA=\frac{SSA}{k-1}=\frac{19.6\text{ 英寸}^2}{3-1}=9.80\text{ 英寸}^2$$

由(18.17)式，有

$$MSW=\frac{SSW}{k(n-1)}=\frac{14.8\text{ 英寸}^2}{3(5-1)}=1.23\text{ 英寸}^2$$

由(18.18)式，有

$$MST=\frac{SST}{kn-1}=\frac{34.4\text{ 英寸}^2}{(3\times5)-1}=2.46\text{ 英寸}^2$$

18.12　F 检验

利用 17.21 节中的方法进行一个 F 检验，即可完成方差分析. 在那一节，我们采用了如下的检验统计量

$$F=\frac{S_1^2}{S_2^2}$$

并且可以根据 F 的具体观测值

$$f = \frac{s_1^2}{s_2^2}$$

来检验零假设 $H_0:\frac{\sigma_1^2}{\sigma_2^2}=1$,其原因在于该统计量服从 F 分布,其分子的自由度为 $v_1=n_1-1$,分母的自由度为 $v_2=n_2-1$,从而保证可以确定临界值决策规则及其 P 值(见 17.21 节).

在方差分析里,现在采用如下检验统计量

$$F=\frac{\dfrac{n\sum_{i=1}^{k}(\bar{X}_{i\cdot}-\bar{X}_{\cdot\cdot})^2}{k-1}}{\dfrac{\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{X}_{i\cdot})^2}{k(n-1)}} \tag{18.21}$$

其观测值为

$$f=\frac{\dfrac{n\sum_{i=1}^{k}(\bar{x}_{i\cdot}-\bar{x}_{\cdot\cdot})^2}{k-1}}{\dfrac{\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_{i\cdot})^2}{k(n-1)}}=\frac{MSA}{MSW} \tag{18.22}$$

从而可检验零假设 $H_0:\mu_1=\mu_2=\cdots=\mu_k$,这是因为该统计量服从 F 分布,分子自由度为 $v_1=k-1$,分母自由度为 $v_2=k(n-1)$,由于 MSA 为 F 的分子,而且我们只关心 MSA 是否大于 MSW,因此该 F 检验的临界值决策规则总为右侧截尾形式[见例 17.11(a)]

$$\text{若 } f^* > f_{\alpha,v_1,v_2}\text{,则拒绝 } H_0 \tag{18.23}$$

其 P 值[(17.50)式]为

$$P=P(F\geqslant f^* \mid H_0 \text{ 为真}) \tag{18.24}$$

式中 F 表示统计量,f_{α,v_1,v_2} 表示 F 临界值,f^* 表示根据样本算得的统计量 F 的一个样本量.

例 18.5 考虑前述的小麦品种试验的例子,采用临界值决策规则,在 $\alpha=0.05$ 时,对 $H_0:\mu_1=\mu_2=\mu_3$ 做右尾检验. 在此以及本章以后的问题里,展示了假设检验的标准步骤(见 16.16 节).

解 ☞ (1) $H_0:\mu_1=\mu_2=\mu_3$,$H_1:\mu_1,\mu_2,\mu_3$ 不全相等;

(2) $\alpha=0.05$;

(3) 假设我们是从具有相同方差的正态总体中独立抽样,则可采用前述的 F 统计量. 利用附表 A.8,相应于 $v_1=k-1=3-1=2$ 与 $v_2=k(n-1)=3(5-1)=12$,其右尾决策规则为:若 $f^*>(f_{0.05,2,12}=3.89)$,则拒绝 H_0;

(4) 统计量 F 的值为

$$f^*=\frac{MSA}{MSW}=\frac{9.80\text{ 英寸}^2}{1.23\text{ 英寸}^2}=7.97$$

由附表 A.8,其 P 值的范围是:$0.005<P<0.01$;

(5) 因 $f^*>3.89$,所以拒绝 H_0 而接受 H_1. 此结果可由 P 值处于范围 $0.05<P<0.01$ 内这一事实得以确认. 因此,在 0.05 水平之下,有显著的证据表明三个均值是不全相同的,即各样本是来自不同的总体.

18.13 方差分析表

表 18.3 称作**方差分析表**或 **ANOVA 表**. 它以标准形式展示了方差分析的结果. 还应记住,由平方和分解式[(18.9)及(18.10)式],我们将"平方和"和"自由度"两栏各取了合计.

表 18.3

方差来源	平方和	自由度	均方	F
组间	SSA	$k-1$	$MSA=\frac{SSA}{k-1}$	$F=\frac{MSA}{MSW}$
组内	SSW	$k(n-1)$	$MSW=\frac{SSW}{k(n-1)}$	
合计	SST	$kn-1$		

例 18.6　用方差分析表给出例 18.5 的方差分析结果

解 ☞　所求方差分析表如表 18.4,其中未标明测量单位.

表 18.4

方差来源	平方和	自由度	均方	F
组间	19.6	2	9.80	7.97
组内	14.8	12	1.23	
合计	34.4	14		

18.14　多重比较检验

在完成单向固定效应方差分析后,我们要么接受 $H_0:\mu_1=\mu_2=\cdots=\mu_k$,要么拒绝 H_0 而接受 $H_1:k$ 个均值不全相同. 若接受 H_0,则整个分析便可结束:没有发现 k 个均值之间存在显著差异. 然而,若接受 H_1,则某些均值之间必然存在较大差异,那么就需做进一步的分析以将这些差异提取出来.

仅简单地对所有可能的配对样本做一系列的双样本检验(见第 17 章)来进行分析是不够的. 其不可取之处就在于:对于多个样本假设 H_0 来说,此种做法增大了犯**第一类错误**的风险(见 16.7 节). 若每个配对比较检验的风险为 α,那么关于 H_0 的多重同时比较检验的风险将大于 α.

为解决此问题,目前已经建立了各种方法,这些方法统称为**多重比较检验**. 对于由方差分析法认定的存在显著差异的均值集合而言,利用这些方法即可把其中相等的均值与不相等的均值分离开来. 这些检验法究竟孰优孰劣,没有统一的看法,在此我们仅选一种被广泛接受的方法来阐明此类检验的步骤. 即 **Duncan 多重极差检验**.

18.15　Duncan 多重极差检验

若方差分析的结果是拒绝 H_0,则接下来可进行 Duncan 多重极差检验. 其各种前提假设与方差分析相同,同时还要求各样本的容量也相同(在 18.23 节我们将给出此检验法在样本容量不同情形下的推广形式). 该检验的核心内容就是:取样本均值 $\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_k$ 的任何 p 个作为一个子集,将此子集的极差与某一统计量的观测值作比较,该值称作"**最小显著极差**",记为 R_p. 若该子集的极差小于 R_p,则认为此子集内的各样本均值是相等的,从而认为相应的总体均值也是相同的. 若该子集的极差大于 R_p,则认为该子集中距离最远的两个样本均值是有显著差异的,从而其相应的总体均值也被认为是不同的.

在操作上,该方法须按一定的标准顺序进行:首先,将最大、最小两个样本均值间的极差与相应的 R_p 值作比较,然后考察最大样本均值与次最小的样本均值间的极差;再考虑最大样本均值与第三小的样本均值间的极差;依此类推,直到最大与次最大的两个样本均值的极差. 接着,要考察从次最大与最小的样本均值的极差,之后是次最大与次最小样本均值的极差,等等. 这种把极差与 R_p 作比较的程序要一直进行下去,直到各样本均值间的关系都得到确定为止.

统计量 R_p 的计算公式为

$$R_p = r_p\sqrt{\frac{MSW}{n}} \tag{18.25}$$

其中数值 r_p 称为**"最小显著学生化极差"**,MSW 是方差分析中的均方,n 是样本容量. 对于子集容量 $p=2,3,\cdots,6$,自由度(关于 MSW 的)$v=1,2,\cdots,20,24,30,40,60,120,\infty$ 及 $\alpha=0.05$与 0.01 的情形,在附录表 A.9(**最小显著学生化极差 r_p**)中给出了相应的 r_p 值.

例 18.7 例 18.5 中,当 $\alpha=0.05$,我们拒绝了 $H_0:\mu_1=\mu_2=\mu_3$. 对于该组数据及同样的显著水平,试利用 Duncan 多重极差检验法判定各均值间的差异.

(在下面的 Duncan 极差检验计算中,没有附带计量单位).

解 (1) 首先,将 $k=3$ 个样本均值按由小到大的顺序排成一行:

$$\begin{matrix} \bar{x}_1 & \bar{x}_2 & \bar{x}_3 \\ 17.2 & 19.4 & 19.8 \end{matrix}$$

(2) 其次,由例 18.4 知 $MSW=1.23$,再由附表 9 中相应于 $\alpha=0.05$, $v=12$ 查到 r_p 的值,如此便可根据(18.25)式计算出 R_p 的值. 所得结果列于下表中:

p	2	3
r_p	3.082	3.225
R_p	1.529	1.600

(3) 在 $p=3$ 时,该子集包含了最大、最小的两个样本均值,其极差为:

$$\bar{x}_3-\bar{x}_1 = 19.8-17.2 = 2.6$$

由于 $2.6>R_3=1.600$,故我们可以认为:$\bar{x}_3$ 与 $\bar{x}_1$ 有显著差异,从而认为:$\mu_3>\mu_1$.

(4) 由于在步骤(3)中,判定了显著差异的存在,则现在必须对该范围内的每对均值做检验. 首先

$$\bar{x}_3-\bar{x}_2 = 19.8-19.4 = 0.4$$

因 $0.4<R_2<1.529$,我们认为 $\bar{x}_3$ 与 $\bar{x}_2$ 差异不显著,从而 $\mu_3=\mu_2$;接下来

$$\bar{x}_2-\bar{x}_1 = 19.4-17.2 = 2.2$$

因 $2.2>R_2=1.529$,于是我们认为 $\bar{x}_2$ 与 $\bar{x}_1$ 有显著差异,从而 $\mu_2>\mu_1$.

(5) 习惯上,我们可以将所有结果按下面方法概括起来:对步骤(1)中的极差,把相邻的无显著差异的 $\bar{x}_i$ 值的下面用横线标为一组:

$$\underline{\bar{x}_1} \quad \underline{\bar{x}_2 \quad \bar{x}_3}$$

由此可下结论:

$$\mu_1<\mu_2=\mu_3$$

18.16 多重比较的后继问题:置信区间的计算

如果方差分析拒绝了 $H_0:\mu_1=\mu_2=\cdots=\mu_k$,并且用多重比较检验,已将 k 个样本均值的差异性区分开来,那么接下来就可以考虑确定个别总体均值及总体均值差的置信区间. 如果某个样本均值 $\bar{x}_i$ 显示出与其他所有样本均值有显著差异,则相应的 μ_i 的$(1-\alpha)100\%$置信区间可由下式确定

$$\bar{x}_i \pm t_{\alpha/2,v}\sqrt{\frac{MSW}{n_i}} \tag{18.26}$$

其中 MSW 是方差分析中的均方,式中的 $n_i=n$ 为方差分析中每个样本的容量,$t_{\alpha/2,v}$中的自

由度 v 为 MSW 的自由度.

如果用多重比较检验法在两个或更多个样本均值间并未发现显著差异,则首先要依下式计算出这些样本均值的总平均

$$\bar{x}_w = \frac{\sum n_i \bar{x}_i}{\sum n_i}$$

这里,参与求和的是被认为不存在显著差异的所有平均值,从而相应的公共均值 μ_w 的 $(1-\alpha)100\%$ 置信区间的计算式为

$$\bar{x}_w \pm t_{\alpha/2,v}\sqrt{\frac{MSW}{\sum n_i}} \tag{18.27}$$

其中的 MSW 与 v 的意义如前,求和过程仍是针对所有等均值的样本.

最后,若发现某两个样本均值 $\bar{x}_i$ 与 $\bar{x}_k$ 间存在显著差异,则可由下式得到 $\mu_i-\mu_k$ 的一个 $(1-\alpha)100\%$ 置信区间

$$(\bar{x}_{i.}-\bar{x}_{k.}) \pm t_{\alpha/2,v}\sqrt{\frac{MSW}{n_i}+\frac{MSW}{n_k}} \tag{18.28}$$

此处的 MSW 及 v 与前面所述相同,n_i 与 n_k 分别是这两个样本的容量. 在等样本容量情形,上式即化作

$$(\bar{x}_i-\bar{x}_k) \pm t_{\alpha/2,v}\sqrt{\frac{2(MSW)}{n}} \tag{18.29}$$

例 18.8　在例 18.7 中,我们曾得到:$\mu_1<\mu_2=\mu_3$. 试由所列数据,计算下面各量的 95% 置信区间:(a) μ_1;(b) μ_2 与 μ_3 的总平均 μ_w;(c) $\mu_2-\mu_1$.

解 ☞　(a) 利用(18.26)式:

$$\bar{x}_i \pm t_{\alpha/2,v}\sqrt{\frac{MSW}{n_i}}$$

且由表 18.2 知 $\bar{x}_1=17.2$ 英寸,$n_i=n=5$. 由附录表 A.6 查得 $t_{\alpha/2,v}=t_{0.05/2,12}=2.179$,再从例 18.4 知 $MSW=1.23$ 英寸^2,所以,μ_1 的 95% 置信区间为

$$17.2\text{英寸} \pm 2.179\sqrt{\frac{1.23\text{英寸}^2}{5}}$$

$$17.2\text{英寸} \pm 1.08\text{英寸}$$

(b) 由(18.27)式,有

$$\bar{x}_w \pm t_{\alpha/2,v}\sqrt{\frac{MSW}{\sum n_i}}$$

其中 $\sum n_i=2\times5=10$,其他各值与(a)中结果相同,且由上册中的(6.19)式有

$$\bar{x}_w=\frac{\sum n_i\bar{x}_i}{\sum n_i}=\frac{(5\times19.4\text{英寸})+(5\times19.8\text{英寸})}{10}=19.6\text{英寸}$$

因而 μ_w 的 95% 置信区间为

$$19.6\text{英寸} \pm 2.179\sqrt{\frac{1.23\text{英寸}^2}{10}}$$

$$19.6\text{英寸} \pm 0.76\text{英寸}$$

(c) 由(18.29)式,有

$$(\bar{x}_i-\bar{x}_k) \pm t_{\alpha/2,v}\sqrt{\frac{2(MSW)}{n}}$$

$$(19.4\text{英寸}-17.2\text{英寸}) \pm 2.179\sqrt{\frac{(2\times1.23\text{英寸}^2)}{5}}$$

2.2英寸 ± 1.53英寸

18.17 方差齐性的检验

在 18.5 节中,我们曾指出,单向固定效应方差分析中有一个关于各方差相等的假定,即认为各总体的方差都等于同一个值 σ^2. 在 17.21 节中,我们对双样本情形,利用方差比较法检验了方差的齐性,可是此方法不能用于多样本场合. 已经有许多相应的多个样本检验法,比如 Cochran检验,最大 F 比检验,还有我们下面要讨论的一种最常用的方法:**Bartlett 检验**.

Bartlett 检验通常有一对要比较的统计假设:零假设

$$H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$$

与备择假设 H_1,一般表作

$$H_1: k \text{ 个方差不全相等}$$

此方法所用的检验统计量为随机变量 B,它近似服从自由度为 $v = k-1$ 的 χ^2 分布(见 15.2 节),只要如方差分析中那样假设 k 个样本来自独立的正态总体即可.

为了确定 B 的观测值 b,从而完成检验过程,首先需对每个样本计算方差 s_i^2,然后用下式将各方差加以平均

$$s_p^2 = \frac{\sum_{i=1}^{k}(n_i - 1)s_i^2}{N - k} \tag{18.30}$$

其中 n_i 表示样本容量,且

$$N = \sum_{i=1}^{k} n_i \tag{18.31}$$

(注:此处 N 用以表示所有样本容量之和,不同于此前的用它来表示总体大小).

接下来就可以依下式计算出 B 的观测值

$$b = 2.3026 \frac{q}{h} \tag{18.32}$$

其中

$$q = (N-k)\log s_p^2 - \sum_{i=1}^{k}(n_i - 1)\log s_i^2 \tag{18.33}$$

$$h = 1 + \frac{1}{3(k-1)}\left(\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{N-k}\right) \tag{18.34}$$

此种形式的 Bartlett 检验使用了常用对数[以 10 为底,见上册的习题 1.23(a)],但也有些写法是采用了自然对数[以 e 为底,见上册的习题 1.23(b)].

量 q 的值当方差均相等时取 0,而当样本方差间差异增大时,其值也随之增大. 因此,这是一个以此为临界值的右尾检验,其决策规则为

$$\text{若 } b^* > \chi^2_{\alpha,v}, \text{则拒绝 } H_0$$

其中 b^* 是 B 的样本观测值,$v = k-1$.

例 18.9 对表 18.2 所给数据,在 $\alpha = 0.05$ 水平下利用 Bartlett 检验法检验方差的齐性.

解 ☞ (省略单位). 由上册的(7.16)式,样本方差为 $s_1^2 = 0.70$, $s_2^2 = 1.30$, $s_3^2 = 1.70$ 利用(18.31)式,对于 $k = 3$,有

$$N = \sum_{i=1}^{3} n_i = 3(5) = 15$$

于是,由(18.30)式,有

$$s_p^2=\frac{\sum_{i=1}^{3}(n_i-1)s_i^2}{15-3}=\frac{14.8}{12}=1.233$$

此外，由(18.33)式，有

$$q=(N-k)\log s_p^2-\sum_{i=1}^{3}(n_i-1)\log s_i^2$$

$$=(12\times0.0910)-0.7580=0.3340$$

再利用(18.34)式得

$$h=1+\frac{1}{6}\left(\sum_{i=1}^{3}\frac{1}{n_i-1}-\frac{1}{15-3}\right)=1.1111$$

则由(18.32)式

$$b^*=2.3026\times\frac{0.3340}{1.1111}=0.692$$

方差的齐性可按如下步骤加以检验：

(1) $H_0:\sigma_1^2=\sigma_2^2=\sigma_3^2$，$H_1$：三个方差不全相等；

(2) $\alpha=0.05$；

(3) 由附录表 A.7，对 $\alpha=0.05$，$v=k-1=3-1=2$，查表得 χ^2 的值，利用 B 统计量，得到临界值决策规则

$$若\ b^*>\chi^2_{0.05,2}=5.99，则拒绝\ H_0$$

(4) B 统计量的样本观测值为 $b^*=0.692$，同时由附录表 A.7 我们看到对于 $v=2$，有一个介于 $\chi^2_{0.900}$ 与 $\chi^2_{0.500}$ 之间的可用以比较的 χ^{2*} 的值，因此 P 值的范围为：$0.5<P<0.9$.

(5) 由于 $b^*<5.99$，我们接受 H_0，即方差是相同的. 这一点可以由 $P>0.05$ 的事实加以确认.

18.18　单向固定效应方差分析：样本容量相等或不等

在 18.6 节至 18.13 节里，在对单向固定效应方差分析探讨之初，我们要求样本容量是相同的 $n_1=n_2=\cdots=n_k=n$. 如今，在 18.18 节至 18.22 节中，我们给出一个处理方差分析的具有一般性的程序，不论样本容量是否相同，它都是适用的.

此时，凡 18.5 节中所涉及的其它假定都保持不变：有一个独立变量与一个因变量，因此水平的选取是出于偏好而不是随机的；相互独立的样本是来自 k 个正态总体，总体均值分别为 $\mu_1,\mu_2,\cdots,\mu_k$，方差分别为 $\sigma_1^2,\sigma_2^2,\cdots,\sigma_k^2$，并且方差是相等的：$\sigma_1^2=\sigma_2^2=\cdots=\sigma_k^2$.

在这个一般的过程中，检验的基本要素是保持不变的，比如：统计假设(见 18.6 节)，数据的整理方式(见 18.7 节)，基本推理(见 18.8 节)，以及各种方法的细节(见 18.9 节至 18.13 节). 当然，为了适应不同的样本容量，各种公式需要有所改变.

18.19　一般程序的单向固定效应方差分析：数据的整理

此处对数据的概括整理方式与表 18.1 是相同的，只不过各描述性统计量的计算发生了变化. 于是，(18.2)，(18.3)，(18.5)式分别变为

$$T_{i.}=\sum_{j=1}^{n_i}x_{ij}\tag{18.35}$$

$$T_{..}=\sum_{i=1}^{k}\sum_{j=1}^{n_i}x_{ij}\tag{18.36}$$

$$\bar{x}_{i.}=\frac{\sum_{j=1}^{n_i}x_{ij}}{n_i}\tag{18.37}$$

此外，利用(18.31)式，则(18.6)式变为

$$T_{..} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}}{N} \tag{18.38}$$

例 18.10 对表 18.2 中的数据,将品种 3 的第 1 个观测值 18 英寸取出,并把它作为品种 1 的第一个观测值,此时,样本容量变为 $n_1=6$, $n_2=5$, $n_3=4$,且由(18.31)式知 $N=\sum_{i=1}^{3} n_i=15$,试将这些数据概括在类似于表 18.1 及表 18.2 的表格之中.

解 ☞ 所求表如表 18.5,其中已给出了合计值与平均值.

表 18.5

小麦品种	高度(英寸)	合计	平均
1	18 17 18 17 16 18	104 英寸	17.3 英寸
2	21 19 18 19 20	97 英寸	19.4 英寸
3	21 20 21 19	81 英寸	20.2 英寸
		282 英寸	18.8 英寸

18.20 一般程序的单向固定效应方差分析:平方和

对于一般程序而言,其平方和的定义式可通过将(18.8),(18.11)与(18.12)式加以修正而得

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_{..})^2 \tag{18.39}$$

$$SSA = \sum_{i=1}^{k} n_i(\bar{x}_{i.}-\bar{x}_{..})^2 \tag{18.40}$$

$$SSW = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_{i.})^2 \tag{18.41}$$

同时 SST[(18.14)式]与 SSA[(18.15)式]的计算公式可相应地改为

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}^2 - \frac{T_{..}^2}{N} \tag{18.42}$$

$$SSA = \sum_{i=1}^{k}\frac{T_{i.}^2}{n_i} - \frac{T_{..}^2}{N} \tag{18.43}$$

而且此时,平方和分解式[(18.9)式]仍然成立,因而可根据(18.13)式计算 SSW: $SSW = SST - SSA$.

例 18.11 对表 18.5 中的数据,利用计算式求 SST、SSA 及 SSW.

解 ☞

$$SST = 5{,}336\text{ 英寸}^2 - (79{,}524\text{ 英寸}^2/15) = 34.40\text{ 英寸}^2$$

$$SSA = 5{,}324.72\text{ 英寸}^2 - (79{,}524\text{ 英寸}^2/15) = 23.12\text{ 英寸}^2$$

$$SSW = 34.40\text{ 英寸}^2 - 23.12\text{ 英寸}^2 = 11.28\text{ 英寸}^2$$

18.21 一般程序的单向固定效应方差分析:自由度与均方

在一般情形下,SSA 的自由度仍为 $k-1$,从而 MSA 的计算式(18.16)变成

$$MSA = \frac{SSA}{k-1} = \frac{\sum_{i=1}^{k} n_i(\bar{x}_{i.}-\bar{x}_{..})^2}{k-1} \tag{18.44}$$

SSW 的自由度为 $N-k$. 为计算 MSW,应将(18.17)式变成

$$MSW=\frac{SSW}{N-k}=\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_{i.})^2}{N-k} \tag{18.45}$$

SST 的自由度为 $N-1$,则(18.18)式应改写为

$$MST=\frac{SST}{N-1}=\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_{..})^2}{N-1} \tag{18.46}$$

例 18.12　利用例 18.11 的结果,计算 MSA 与 MSW.

解 ☞

$$MSA=\frac{23.12\text{英寸}^2}{2}=11.56\text{英寸}^2$$

$$MSW=\frac{11.28\text{英寸}^2}{12}=0.94\text{英寸}^2$$

18.22　一般程序的单向固定效应方差分析:F 检验

此时,用以检验 $H_0:\mu_1=\mu_2=\cdots=\mu_k$ 的检验统计量是将(18.21)式变为

$$F=\frac{\dfrac{\sum_{i=1}^{k}n_i(\bar{X}_{i.}-\bar{X}_{..})^2}{k-1}}{\dfrac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_{i.})^2}{N-k}} \tag{18.47}$$

由此可计算其样本观测值

$$f=\frac{\dfrac{\sum_{i=1}^{k}n_i(\bar{x}_{i.}-\bar{x}_{..})^2}{k-1}}{\dfrac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_{i.})^2}{N-k}}=\frac{MSA}{MSW} \tag{18.48}$$

而检验的临界值决策规则[(18.23)式]中,取 $v_1=k-1$, $v_2=N-k$;P 值的计算式[(18.24)式]与前面的相同.

例 18.13　对表 18.5 所列数据,并利用例 18.11 及例 18.12 的结果,试在 $\alpha=0.05$ 水平下,利用临界值决策规则,对 $H_0:\mu_1=\mu_2=\mu_3$ 进行右尾检验.

解 ☞　(1) $H_0:\mu_1=\mu_2=\mu_3$,H_1:三个均值不全相等;

(2) $\alpha=0.05$;

(3) 此处右尾决策规则与例 18.5(3)的形式相同

$$\text{若 } f^*>f_{0.05,2,12}=3.89\text{,则拒绝 } H_0$$

(4) 方差分析的结果显示于下面的表 18.6 中. 由附表 A.8 可知 P 值的范围为 $P<0.005$.

表 18.6

方差来源	平方和	自由度	均方	F
组间	23.1	2	11.56	12.30
组内	11.28	12	0.94	
合计	34.40	14		

(5) 因为 $f^* > 3.89$,所以拒绝 H_0 而接受 H_1,此结果可以由 $P<0.005$ 的事实得以确认. 从而在 0.05 的显著水平之下,有证据表明三个均值是不全相等的.

18.23 一般程序的单向固定效应方差分析:多重比较

在 18.15 节中所采用的 Duncan 多重极差检验是多重比较检验的一种,它要求样本容量彼此相同. 不过,能处理一般情形的多重比较检验也是有的,它对样本容量相等或不相等二者都是适用的. 我们引入一种由 C. Y. Kramer 建立的 Duncan 检验法的推广形式.

这种检验法也是从样本均值 $\bar{x}_1, \cdots, \bar{x}_k$ 任取 p 个生成的子集的极差与一个最小显著极差进行比较. 此处,该极差的值用 R'_p 表示,其计算公式为

$$R'_p = \sqrt{MSW} r_p \tag{18.49}$$

其中 r_p 的值(最小显著学生化极差)也可以在附表 A.9 中通过取 $v = N - k$ 而得到. 式中的 MSW 是方差分析中的均方. 不过,为了能够适应样本容量不相等的情况,必须把用来与 R'_p 作比较的平均值的极差修正为

$$(\bar{x}_i - \bar{x}_k)\sqrt{\frac{2n_i n_k}{n_i + n_k}} \tag{18.50}$$

其中 $\bar{x}_i$. 是两个平均值中较大者.

例 18.14 在例 18.13 中,当 $\alpha = 0.05$ 时,我们拒绝了 $H_0: \mu_1 = \mu_2 = \mu_3$,利用同一数据(见表 18.5),在同一显著性水平之下,采用由 C. Y. Kramer 推广的 Duncan 多重比较检验法来确定各均值间的差异.

解 ☞ (1) 将 $k = 3$ 个样本均值从小到大排成一行:

$$\begin{matrix} \bar{x}_1 & \bar{x}_2 & \bar{x}_3 \\ 17.3 & 19.4 & 20.2 \end{matrix}$$

(2) 由例 18.12 知 $MSW = 0.94$,给定 $\alpha = 0.05$, $v = N - k = 12$,查附表 A.9 得 r_p 值,从而根据(18.49)式可算出 R'_p,计算结果见下表

p	2	3
r_p	3.082	3.225
R'_p	2.988	3.127

(3) 对 $p = 3$ 的子集.它包含了最大样本均值与最小样本均值,利用(18.50)式,其极差为

$$(\bar{x}_3 - \bar{x}_1)\sqrt{\frac{2n_3 n_1}{n_3 + n_1}} = (20.2 - 17.3)\sqrt{\frac{48}{10}} = 6.354$$

由于 $6.354 > R'_p = 3.127$,故认为 $\bar{x}_3$ 与 $\bar{x}_1$ 显著不同,所以 $\mu_3 > \mu_1$.

(4) 由于在步骤(3)中确定了判差异性的存在,故需在该范围内进行两两配对检验. 首先,

$$(\bar{x}_3 - \bar{x}_2)\sqrt{\frac{2n_3 n_2}{n_3 + n_2}} = (20.2 - 19.4)\sqrt{\frac{40}{9}} = 1.687$$

则由于 $1.687 < R'_2 = 2.988$,故可认为 $\bar{x}_3$ 与 $\bar{x}_2$ 无显著差异,从而 $\mu_3 = \mu_2$.

其次,

$$(\bar{x}_2 - \bar{x}_1)\sqrt{\frac{2n_2 n_1}{n_2 + n_1}} = (19.4 - 17.3)\sqrt{\frac{60}{11}} = 4.905$$

则由于 $4.905 > R'_2 = 2.988$,故可以认为 $\bar{x}_2$ 与 $\bar{x}_1$ 有显著差异,从而 $\mu_2 > \mu_1$.

(5) 由以上结果,可做出结论

$$\underline{\bar{x}_1} \quad \underline{\bar{x}_2 \quad \bar{x}_3}$$

因此我们认为 $\mu_1<\mu_2=\mu_3$.

18.24　一般程序的单向固定效应方差分析:置信区间的计算及方差齐性的检验

在 18.16 节中,在多重比较之后,(18.26),(18.27)与(18.28)式给出了计算置信区间的一般方法(样本容量相等或不等). 同时在 18.17 节中还给出了多个样本场合中关于方差齐性检验的一般方法.

例 18.15　我们在例 18.14 中得到 $\mu_1<\mu_2=\mu_3$,利用该组数据,求 $\mu_2-\mu_1$ 的 95%置信区间.

解 ☞　根据(18.28)式及表 18.5 与表 18.6 中的一些结果,并利用附表 A.6,有

$$(\bar{x}_2-\bar{x}_1)\pm t_{0.05/2,12}\sqrt{\frac{MSW}{n_2}+\frac{MSW}{n_1}}$$

$$(19.4\text{ 英寸}-17.3\text{ 英寸})\pm 2.179\sqrt{\frac{0.94\text{ 英寸}^2}{5}+\frac{0.94\text{ 英寸}^2}{6}}$$

$$2.1\text{ 英寸}\pm 12.28\text{ 英寸}$$

例 18.16　对表 18.5 中的数据,在 $\alpha=0.05$ 的水平下,利用 Bartlett 检验法对方差的齐性作检验.

解 ☞　利用上册中的(7.16)式可知,样本方差为 $s_1^2=0.668$, $s_2^2=1.300$、$s_3^2=0.920$. 对于 $k=3$,利用(18.31)式,有

$$N=\sum_{i=1}^{3}n_i=6+5+4=15$$

再利用(18.30)式,有

$$s_p^2=\frac{\sum_{i=1}^{3}(n_i-1)s_i^2}{N-k}=\frac{11.30}{12}=0.942$$

从而由(18.33)式,得

$$\begin{aligned}q&=(N-k)\log s_p^2-\sum_{i=1}^{3}(n_i-1)\log s_i^2\\&=[12\times(-0.0259)]-(-0.5290)=0.2182\end{aligned}$$

又由(18.34)式,有

$$h=1+\frac{1}{6}\left(\sum_{i=1}^{3}\frac{1}{n_i-1}-\frac{1}{15-3}\right)=1.1167$$

故由(18.32)式,得

$$b^*=2.3026\times\frac{0.2182}{1.1167}=0.450$$

关于方差齐性的检验总结如下:

(1) $H_0:\sigma_1^2=\sigma_2^2=\sigma_3^2$, H_1:三个方差不全相等;

(2) $\alpha=0.05$;

(3) 对于 $\alpha=0.05$, $v=k-1=3-1=2$,查表 A.7 得 B 统计量的临界值,从而,临界值决策规则为:若 $b^*>\chi^2_{0.05,2}=5.99$,则拒绝 H_0;

(4) 已知 B 统计量的样本值 $b^*=0.450$,由表 A.7 可得位于 $\chi^2_{0.900}$ 与 $\chi^2_{0.500}$ 之间的一个 χ^{2*} 值,相应 P 值为:$0.5<P<0.9$;

(5) 由于 $b^*<5.99$,故接受 H_0,即方差存在齐性,这一结论可由事实:$P>0.05$ 确认.

18.25　方差分析前提假定的破坏

单向固定效应方差分析作为多个样本与单样本(见 16.24 节)及双样本(见 17.13 节)方法一样,也被称为参数方法,其原因在于:它们都用于对参数进行估计和假设检验,并且许多前提假定也是关于参数的,因此,方差分析中的假定(见 18.5 节)也被称为参数假设.

在方差分析中,有一个基本的抽样假设必须满足:对取间隔尺度或比例尺度进行独立随机

抽样. 然而,像 t 检验法(见 16.25 节,17.14 节)那样,在正态分布的假定之下. 方差分析法是相当稳健的,尤其对于大样本更是如此,而且在方差齐性的假定下,它还有适度的稳健性,尤其当各样本容量相等时:$n_1 = n_2 = \cdots = n_k = n$.

过份偏离正态假定与方差齐性假定可能会严重影响方差分析的效果. 所以在单样本及双样本方法(见 16.25 节与 17.14 节)里,第一步就是分析抽样数据以查看前提假定是否得到了很好的满足. 在 16.25 节中,我们给出了对正态分布假定加以评估的方法,在 18.17 节与 18.24 节中,我们给出了在方差分析里如何去检验方差的齐性. 不过,在前提假定遭到破坏时,单向固定效应方差分析法比 t 检验法(见 17.14 节)更加稳健,从而它是齐性检验. 据此,特别在样本容量相同的情形下,我们在试图进行方差分析之前,对于方差齐性的检验可以不作要求.

如果对于前提假定的严重偏离已得到确认,则下一步应考虑对数据做一个变换. 正像在单样本(见 16.25 节)与双样本(见 17.14 节)方法中要求的那样,在方差分析里,也需要同样的变换. 若变换是成功的,则就可以对变换后的数据进行方差分析.

再者,仍类似于单样本与双样本方法情形,如果找不到使假定得以满足的变换,那么,可以采用非参数方法(见第 20 章)对数据从其它方面进行分析.

习 题 解 答

样本容量相等时的单向固定效应方差分析:H_0 与 H_1

18.1 某运动生理学家提出了一个假说:对于耐力型运动员,比如长跑运动员,如果他们在高海拔处睡觉而在低海拔处训练的话,则他们在比赛中会发挥得最出色. 为检证此假说,他对 20 名从事 10 千米赛跑的运动员进行测试,这些运动员以前都一直在海平面高处生活和训练. 他们被随机地平分为四个组:(1)SH/TL(在 8,200 英尺高处的山城睡觉,每天在 4,500 英尺高处训练);(2)SH/TH(在 8,200 英尺高处睡觉和训练);(3)SL/TH(在 4,500 英尺高处睡觉,在 8,200 英尺高处训练);(4)SL/TL(在 4,500 英尺高处睡觉和训练). 在某个夏天,对运动员们按同样常规进行了两个月的训练. 在训练之初,所有人都在 4,500 英尺高处进行了 10 千米赛跑,各个人所用时间均被记录下来. 经过了两个月训练后,再于 4,500 英尺高处测试并记录了他们每人的 10 千米跑成绩. 所要分析的数据即是每个人的先后两个成绩之差距(以秒计):从第一次成绩中减去第二次成绩. 假定所有的差距的总体服从正态分布,其均值分别为 $\mu_1, \mu_2, \mu_3, \mu_4$,有公共方差:$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$. 若该生理学家准备对此组数据做一个单向固定效应方差分析,问:什么是 H_0,什么是 H_1?

解 ☞ $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$,$H_1$:四个均值不全相等.

18.2 某制造商生产某种以电池供电的手提电话,他正考虑在某种新电话上使用三种电池(类型 1,2 与 3). 假设你被要求判断这三种电池的通话时间(即使用寿命)是否存在差异. 你在每种电池中各随机选取 10 节,然后测试每节电池在新电话上连续使用的通话时间(精确到 0.1 小时). 假设这些时间的总体服从正态分布,分别有均值 μ_1, μ_2, μ_3,并且有公共方差:$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$. 若你打算对这些数据做一个单向固定效应方差分析,则 H_0 与 H_1 怎样表述?

解 ☞ $H_0: \mu_1 = \mu_2 = \mu_3$,$H_1$:三个均值不全相等.

样本容量相等时的方差分析:整理数据

18.3 对于习题 18.1 中的训练试验,现在给出每个人的时间差(精确到秒):第 1 组(11,13,10,12,11);第 2 组(8,9,11,10,9);第 3 组(7,8,10,9,7);第 4 组(6,7,8,6,8). 试将这些数据整理成形如表 18.1 那样的表.

解 ☞　所求表即后面的表 18.7,其间已根据(18.2),(18.3)及(18.5)式给出了合计值与平均值.

表 18.7

睡眠/训练组	时间差(秒)					合计	平均
1	11	13	10	12	11	57 秒	11.4 秒
2	8	9	11	10	9	47 秒	9.4 秒
3	7	8	10	9	7	41 秒	8.2 秒
4	6	7	8	6	8	35 秒	7.0 秒
						180 秒	9.0 秒

18.4　对于习题 18.2 中所述的电池试验,下面是每组的平均通话时间:$\bar{x}_1=1.34$ 小时,$\bar{x}_2=1.26$ 小时,$\bar{x}_3=1.06$ 小时,试确定 T_1,T_2,T_3,T 与 $\bar{\bar{x}}$ 的值.

解 ☞　利用(18.2)与(18.5)式,得

$$T_1=n\times\bar{x}_1=10\times1.34\text{ 小时}=13.4\text{ 小时}$$

$$T_2=n\times\bar{x}_2=10\times1.26\text{ 小时}=12.6\text{ 小时}$$

$$T_3=n\times\bar{x}_3=10\times1.06\text{ 小时}=10.6\text{ 小时}$$

利用(18.2)与(18.3)式,得

$$T=T_1+T_2+T_3=13.4\text{ 小时}+12.6\text{ 小时}+10.6\text{ 小时}=36.6\text{ 小时}$$

由(18.6)式,得

$$\bar{\bar{x}}=\frac{T}{kn}=\frac{36.6\text{ 小时}}{10\times3}=1.22\text{ 小时}$$

样本容量相等时的方差分析:平方和

18.5　针对以上的训练试验,试对表 18.7 中的数据,利用(18.13),(18.14)与(18.15)式计算 SST,SSA 与 SSW.

解 ☞

$$SST=\sum_{i=1}^{4}\sum_{j=1}^{5}x_{ij}^2-\frac{T^2}{kn}=1{,}694\text{ 秒}^2-(32{,}400\text{ 秒}^2/20)=74.0\text{ 秒}^2$$

$$SSA=\frac{\sum_{i=1}^{4}T_i^2}{n}-\frac{T^2}{kn}=(8{,}364\text{ 秒}^2/5)-(32{,}400\text{ 秒}^2/20)=52.8\text{ 秒}^2$$

$$SSW=SST-SSA=74.0\text{ 秒}^2-52.8\text{ 秒}^2=21.2\text{ 秒}^2$$

18.6　针对上述的电池试验,根据习题 18.4 的结果及平方和 $\sum_{i=1}^{3}\sum_{j=1}^{10}x_{ij}^2=45.62$ 小时^2,并利用(18.13),(18.14)与(18.15)式计算 SST,SSA 及 SSW.

解 ☞

$$SST=\sum_{i=1}^{3}\sum_{j=1}^{10}x_{ij}^2-\frac{T^2}{kn}=45.62\text{ 小时}^2-(1{,}339.56\text{ 小时}^2/30)=0.968\text{ 小时}^2$$

$$SSA=\frac{\sum_{i=1}^{3}T_i^2}{n}-\frac{T^2}{kn}=(450.68\text{ 小时}^2/10)-(1{,}339.56\text{ 小时}^2/30)=0.416\text{ 小时}^2$$

$$SSW=SST-SSA=0.968\text{ 小时}^2-0.416\text{ 小时}^2=0.552\text{ 小时}^2$$

样本容量相等时的方差分析:均方

18.7　针对前述的训练试验,利用习题 18.5 中 SSA 与 SSW 的结果,计算 MSA 与 MSW.

解 ☞　由(18.16)式,得

$$MSA=\frac{SSA}{k-1}=\frac{52.8\text{秒}^2}{4-1}=17.60\text{秒}^2$$

由(18.17)式,得

$$MSW=\frac{SSW}{k(n-1)}=\frac{21.2\text{秒}^2}{4(5-1)}=1.32\text{秒}^2$$

18.8 针对电池试验,利用习题 18.6 中 SSA 与 SSW 的值,计算 MSA 与 MSW.

解 ☞ 由(18.16)式,得

$$MSA=\frac{SSA}{k-1}=\frac{0.416\text{小时}^2}{3-1}=0.2080\text{小时}^2$$

由(18.17)式,得

$$MSW=\frac{SSW}{k(n-1)}=\frac{0.552\text{小时}^2}{3(10-1)}=0.0204\text{小时}^2$$

样本容量相等时的方差分析:F 检验

18.9 考虑前述的训练试验,利用习题 18.1,18.3,18.5,18.7 的结果,在 $\alpha=0.05$ 的显著性水平下,采用临界值决策规则,对零假设 $H_0:\mu_1=\mu_2=\mu_3=\mu_4$,做右尾试验.

解 ☞ (1) $H_0:\mu_1=\mu_2=\mu_3=\mu_4$, H_1:四个均值不全相等;

(2) $\alpha=0.05$;

(3) 由于我们假设是从具有相同方差的正态总体中进行独立抽样,所以可采用 F 统计量[(18.21)式]. 故由附录表 A.8[$v_1=k-1=4-1=3$, $v_2=k(n-1)=4(5-1)=16$]可得右尾决策规则:若 $f^*>(f_{0.05,3,16}=3.24)$,则拒绝 H_0

(4) 方差分析的结果见表 18.8,由附录表 A.8 可知 P 值的范围为 $P<0.005$.

表 18.8

方差来源	平方和	自由度	均方	F
组间	52.8	3	17.60	3.33
组内	21.2	16	1.32	
合计	74.0	19		

(5) 由于 $f^*>3.24$,故拒绝 H_0,接受 H_1,此结论可由 $P<0.005$ 这一事实加以确认. 从而在 0.05的显著水平下,有证据表明四个均值不全相等.

18.10 考虑电池试验,利用习题 18.2,18.4,18.6,18.8 所给信息,在 $\alpha=0.05$ 的显著性水平下,采用临界值决策规则,对零假设 $H_0:\mu_1=\mu_2=\mu_3=\mu_4$,做右尾试验.

解 ☞ (1) $H_0:\mu_1=\mu_2=\mu_3=\mu_4$, H_1:四个均值不全相等;

(2) $\alpha=0.05$;

(3) 由于我们假设是从具有相同方差的正态总体中进行独立抽样,所以可采用 F 统计量[(18.21)式]. 故由附录表 8[$v_1=k-1=3-1=2$, $v_2=k(n-1)=3(10-1)=27$]可得右尾决策规则:若 $f^*>(f_{0.05,2,27}=3.35)$,则拒绝 H_0;

(4) 方差分析的结果见表 18.9,由附录表 A.8 可知 P 值的范围为 $P<0.005$;

表 18.9

方差来源	平方和	自由度	均方	F
组间	0.416	2	0.2080	10.20
组内	0.552	27	0.0204	
合计	0.968	29		

(5) 由于 $f^*>3.35$,故拒绝 H_0,接受 H_1,此结论可由 $P<0.005$ 这一事实加以确认. 从而在 0.05的显著水平下,有证据表明四个均值不全相等.

样本容量相等时的方差分析:多重比较

18.11　在习题 18.9 中,对于训练试验,我们在 $\alpha=0.05$ 时,拒绝了零假设 $H_0:\mu_1=\mu_2=\mu_3=\mu_4$,试在同一显著性水平下,利用这些数据,采用 Duncan 多重极差检验(见 18.15 节),确定这些均值间的差异.

解 ☞　(1) 首先,将 4 个样本均值从小到大排成一行

$$\begin{matrix}\bar{x}_4 & \bar{x}_3 & \bar{x}_2 & \bar{x}_1\\ 7.0 & 8.2 & 9.4 & 11.4\end{matrix}$$

(2) 由习题 18.7,$MSW=1.32$,取 $\alpha=0.05$ 及 $v=16$,在附录表 A.9 中查到 r_p 的值,然后用(18.25)式算得 R_p 的值,结果如下

p	2	3	4
r_p	2.988	3.144	3.235
R_p	1.540	1.615	1.662

(3) 当 $p=4$ 时,相应子集合有最大样本均值与最小样本均值,其极差为

$$\bar{x}_1-\bar{x}_4=11.4-7.0=4.4$$

由于 $4.4>R_4=1.662$,我们认为 $\bar{x}_1$ 与 $\bar{x}_4$ 是显著不同的,从而 $\mu_1>\mu_4$.

(4) 由于在步骤(3)中发现了显著差异,则须检查 $\bar{x}_1$ 与 $\bar{x}_3$ 的差异

$$\bar{x}_1-\bar{x}_3=11.4-8.2-3.2$$

因 $3.2>R_3=1.615$,则认为 $\bar{x}_1$ 与 $\bar{x}_3$ 间是有显著差异的,从而 $\mu_1>\mu_3$.

(5) 由于在步骤(4)中存在显著差异,则还须考察 $\bar{x}_1$ 与 $\bar{x}_2$ 的差异

$$\bar{x}_1-\bar{x}_2=11.4-9.4=2.0$$

因 $2.0>R_2=1.540$,则认为 $\bar{x}_1$ 与 $\bar{x}_2$ 差异显著,从而 $\mu_1>\mu_2$.

(6) 接下来,考察次最大样本均值与最小样本均值

$$\bar{x}_2-\bar{x}_4=9.4-7.0=2.4$$

由于 $2.4>R_3=1.615$,则认为 $\bar{x}_2$ 与 $\bar{x}_4$ 显著差异,从而 $\mu_2>\mu_4$.

(7) 由于在步骤(6)中存在显著差异,今须检验 $\bar{x}_2$ 与 $\bar{x}_3$

$$\bar{x}_2-\bar{x}_3=9.4-8.2=1.2$$

因 $1.2<R_2=1.540$,则认为 $\bar{x}_2$ 与 $\bar{x}_3$ 无显著差异,从而 $\mu_2=\mu_3$.

(8) 最后,我们检验 $\bar{x}_3$ 与 $\bar{x}_4$ 的差异

$$\bar{x}_3-\bar{x}_4=8.2-7.0=1.2$$

由于 $1.2<R_2=1.540$,则认为 $\bar{x}_3$ 与 $\bar{x}_4$ 无显著差异,从而 $\mu_3-\mu_4$.

(9) 由上述可知

$$\bar{x}_4\quad \underline{\bar{x}_3\quad \bar{x}_2}\quad \underline{\bar{x}_1}$$

故此,我们可以明确地下结论:$\mu_4<\mu_2<\mu_1$,但是,由这个检验,我们尚不能判断 μ_3 与其他值的关系如何.

18.12　在习题 18.10 中,对于前述电池试验,我们在 $\alpha-0.05$ 时拒绝了 $H_0:\mu_1=\mu_2=\mu_3$. 针对该组数据,试在同样的显著水平下,利用 Duncan 多重极差检验(见 18.15 节)确定这些均值间的差异.

解 ☞　(1) 首先,将 3 个样本均值从小到大排成一行

$$\bar{x}_3\quad \bar{x}_2\quad \bar{x}_1$$

1.06 1.26 1.34

(2) 其次,由习题 18.8 知 $MSW=0.0204$,我们试图针对 $\alpha=0.05$ 与 $v=27$ 从附录表 A.9 查得 r_p 的值并进而求出 R_p 的值. 可是,附录表 A.9 仅针对 $v=24$ 及 $v=30$ 的情形给出了 r_p 的值,则必须利用线性插值方法(见例 14.15)获取 $v=27$ 时的 r_p 值. 这样,对于 $p=2$,我们在 $v=24$ 的 $r_2=2.919$ 与 $v=30$ 的 $r_2=2.888$ 间进行插值,即可获得 $v=27$ 时相应的 $r_2=2.904$. 对于 $p=3$,在 $v=24$ 的 $r_3=3.066$ 与 $v=30$ 的 $r_3=3.035$ 间进行插值,即得到 $v=27$ 时,相应 $r_3=3.050$. 利用这些结果,便可由(18.25)式算得 R_p 值,各值结果见下表

p	2	3
r_p	2.904	3.050
R_p	0.131	0.138

(3) 对 $p=3$ 情形,相应子集包含了最大样本均值与最小样本均值,其极差为

$$\bar{x}_1-\bar{x}_3=1.34-1.06=0.28$$

则由于 $0.28>R_3=0.138$,我们认为 $\bar{x}_1$ 与 $\bar{x}_3$ 有显著差异,从而 $\mu_1>\mu_3$.

(4) 因为在步骤(3)中存在显著差异,则必须对此极差进行两两检验. 首先,

$$\bar{x}_1-\bar{x}_2=1.34-1.26=0.08$$

则由于 $0.08<R_2=0.131$,我们认为 $\bar{x}_1$ 与 $\bar{x}_2$ 无显著差异,从而 $\mu_1=\mu_2$. 然后,我们检验

$$\bar{x}_2-\bar{x}_3=1.26-1.06=0.20$$

则由于 $0.20>R_2=0.131$,我们认为 $\bar{x}_2$ 与 $\bar{x}_3$ 有显著差异,从而 $\mu_2>\mu_3$.

(5) 由上述所有结果,可下结论

$$\underline{\bar{x}_3}\quad \underline{\bar{x}_2\quad \bar{x}_1}$$

因此可知 $\mu_3<\mu_2=\mu_1$.

样本容量相等时的方差分析:置信区间

18.13 在习题 18.11 中,我们确定了 $\mu_4<\mu_2<\mu_1$. 试根据所给数据计算 μ_1 的 95% 置信区间.

解 由(18.26)式有

$$\bar{x}_i \pm t_{\alpha/2,v}\sqrt{\frac{MSW}{n_i}}$$

根据表 18.7 可知 $\bar{x}_1=11.4$ 秒,$n_1=5$. 从附录表 A.6 查得 $t_{\alpha/2,v}=t_{0.05/2,16}=2.120$,再从习题 18.7 获知 $MSW=1.32$ 秒^2. 则 μ_1 的 95% 置信区间为

$$11.4\text{秒}\pm 2.120\sqrt{\frac{1.32\text{秒}^2}{5}}$$

即 11.4 秒 ± 1.09 秒.

18.14 在习题 18.12 中已确认 $\mu_3<\mu_2=\mu_1$. 试根据其数据计算 μ_1 与 μ_2 的总均值 μ_u 的一个 95% 置信区间.

解 由(18.27)式有

$$\bar{x}_u \pm t_{\alpha/2,v}\sqrt{\frac{MSW}{\sum n_i}}$$

其中 $\sum n_i=2\times 10=20$,利用上册中的(6.19)式及习题 18.4 的结果可知

$$\bar{x}_w=\frac{\sum n_i\bar{x}_i}{\sum n_i}=\frac{(10\times 1.34\text{小时})+(10\times 1.26\text{小时})}{20}=\frac{26.0\text{小时}}{20}=1.30\text{小时}$$

从习题 18.8 知 $MSW=0.0204$ 小时^2,再由附录表 A.6 知 $t_{\alpha/2,v}=t_{0.05/2,27}=2.052$. 则 μ_u 的一个

95%置信区间为

$$1.30\text{小时} \pm 2.052\sqrt{\frac{0.0204\text{小时}^2}{20}}$$

即

$$1.30\text{小时} \pm 0.066\text{小时}$$

样本容量相等时的方差分析:方差齐性的检验

18.15 根据表 18.7 所给数据,在 $\alpha=0.05$ 水平下用 Bartlett 检验法检验方差的齐性.

解 由上册中的(7.16)式,样本方差为 $s_1^2=1.30, s_2^2=1.30, s_3^2=1.70, s_4^2=1.00$,对 $k=4$ 利用(18.31)式有

$$N=\sum_{i=1}^{4} n_i = 4(5)=20$$

然后,由(18.30)式知

$$s_p^2=\frac{\sum_{i=1}^{4}(n_i-1)s_i^2}{N-k}=\frac{21.20}{16}=1.325$$

再由(18.33)式有

$$\begin{aligned} q &= (N-k)\log s_p^2-\sum_{i=1}^{4}(n_i-1)\log s_i^2 \\ &= (16\times 0.1222)-1.8333=0.1219 \end{aligned}$$

接着,由(18.34)式知

$$h=1+\frac{1}{9}\left(\sum_{i=1}^{4}\frac{1}{n_i-1}-\frac{1}{20-4}\right)=1.1042$$

则根据(18.32)式得

$$b^*=2.3026\times\frac{0.1219}{1.1042}=0.254$$

方差齐性检验可概括如下:

(1) $H_0:\sigma_1^2=\sigma_2^2=\sigma_3^2=\sigma_4^2$, H_1:四个方差不全相等;

(2) $\alpha=0.05$;

(3) 对于 $\alpha=0.05$, $v=k-1=4-1=3$,由附录表 A.7,利用 B 统计量,知临界值决策规则为:若 $b^*>(\chi^2_{0.05,3}=7.81)$,则拒绝 H_0;

(4) B 统计量的样本观测值为 $b^*=0.254$,且由附录表 A.7 可见对于 $v=3$,存在一个介于 $\chi^2_{0.975}$ 与 $\chi^2_{0.950}$ 之间的可供比较的 χ^{2*} 的值,因而 P 值范围是 $0.950<P<0.975$;

(5) 因 $b^*<7.81$,故接受 H_0,即方差是相等的.此结论可由事实 $P>0.05$ 加以确认.

18.16 对于习题 18.2 中的电池试验,其样本方差为 $s_1^2=0.01822, s_2^2=0.02267, s_3^2=0.02044$,由此结果,在 $\alpha=0.05$ 水平下,利用 Bartlett 方法检验方差的齐性.

解 由 $k=3$,利用(18.31)式有

$$N=\sum_{i=1}^{3} n_i=3(10)=30$$

则由(18.30)式知

$$s_p^2=\frac{\sum_{i=1}^{3}(n_i-1)s_i^2}{N-k}=\frac{0.55197}{27}=0.02044$$

再由(18.33)式得

$$q=(N-k)\log s_p^2-\sum_{i=1}^{3}(n_i-1)\log s_i^2$$

$$=[27\times(-1.6895)]-(-45.6617)=0.04520$$

然后,利用(18.34)式知

$$h=1+\frac{1}{6}\left(\sum_{i=1}^{3}\frac{1}{n_i-1}-\frac{1}{20-3}\right)=1.04575$$

故由(18.32)式有

$$b^*=2.3026\times\frac{0.04520}{1.04575}=0.0995$$

方差齐性检验可概括如下:

(1) $H_0:\sigma_1^2=\sigma_2^2=\sigma_3^2$, H_1:三个方差不全相等;

(2) $\alpha=0.05$;

(3) 对 $\alpha=0.05$, $v=k-1=3-1=2$,由附录表 A.7 利用 B 统计量可知临界值决策规则为:若 $b^*>(\chi^2_{0.05,2}=5.99)$,则拒绝 H_0;

(4) B 统计量的样本观测值为 $b^*=0.0995$,且由附录表 A.7 可见对于 $v=2$,有一个介于 $\chi^2_{0.975}$ 与 $\chi^2_{0.950}$ 间的可供比较的 χ^{2*} 的值,从而 P 值的范围是 $0.950<P<0.975$;

(5) 由于 $b^*<5.99$,故接受 H_0,即方差是相等的. 此结论可由 $P>0.05$ 的事实加以确认.

一般程序的单向固定效应方差分析:H_0 与 H_1

18.17 再回到习题 18.2 的电池试验,仍然是在三个品类中各取 10 节在新电话上连续使用并记录其通话时间. 这次,由于电话上的技术失误导致某些电池未能完成试验:第 1 型电池 2 节,第 3 型电池 1 节,余下 27 节电池的试验结果如下(精确到 0.1 小时):组 1(1.5,1.1,1.3,1.4,1.5,1.4,1.2,1.4);组 2(1.4,1.0,1.2,1.3,1.2,1.0,1.3,1.2,1.4,1.3);组 3(1.0,1.1,1.3,0.9,1.2,1.0,1.1,1.2,1.0). 仍然假定各时间总体服从正态分布,其均值为 μ_1,μ_2,μ_3,方差彼此相等:$\sigma_1^2=\sigma_2^2=\sigma_3^2=\sigma^2$. 如果要对此数据做一个单向固定效应方差分析,则何为 H_0 与 H_1?

解 ☞ $H_0:\mu_1=\mu_2=\mu_3$, $H_1:\mu_1,\mu_2,\mu_3$ 不全相等.

一般程序的方差分析:整理数据

18.18 对习题 18.17 中的电池试验,试将其数据整理成为如表 18.1 那样的表.

解 ☞ 下面的表 18.10 即为所求,其中已根据(18.35),(18.36)及(18.38)式算出了合计值与平均值.

表 18.10

电池类型	时间(小时)	合计	平均
1	1.5 1.1 1.3 1.4 1.5 1.4 1.2 1.4	10.8 小时	1.35 小时
2	1.4 1.0 1.2 1.3 1.2 1.0 1.3 1.2 1.4 1.3	12.3 小时	1.23 小时
3	1.0 1.1 1.3 0.9 1.2 1.0 1.1 1.2 1.0	9.8 小时	1.09 小时
		32.9 小时	1.22 小时

一般程序的方差分析:平方和

18.19 针对表 18.10 所列数据,利用计算式求 SST, SSA 与 SSW.

解 ☞ 由(18.31)式,

$$N=\sum_{i=1}^{3}n_i=8+10+9=27$$

由(18.42)式,有

$$SST=\sum_{i}^{3}\sum_{j=1}^{n_i}x_{ij}^2-\frac{T^2}{N}=40.83\text{ 小时}^2-(1,082.41\text{ 小时}^2/27)=0.741\text{ 小时}^2$$

由(18.43)式,有

$$SSA=\sum_{i=1}^{3}\frac{T_{i\cdot}^2}{n_i}-\frac{T^2}{N}=40.38\text{ 小时}^2-(1{,}082.41\text{ 小时}^2/27)=0.291\text{ 小时}^2$$

由(18.9)式，有

$$SSW=SST-SSA=0.741\text{ 小时}^2-0.291\text{ 小时}^2=0.450\text{ 小时}^2$$

一般程序的方差分析：均方

18.20 根据习题 18.19 的结果，计算 MSA 与 MSW.

解 ☞ 由(18.44)式，有

$$MSA=\frac{SSA}{k-1}=\frac{0.291\text{ 小时}^2}{3-1}=0.146\text{ 小时}^2$$

由(18.45)式，有

$$MSW=\frac{SSW}{N-k}=\frac{0.450\text{ 小时}^2}{27-3}=0.019\text{ 小时}^2$$

一般程序下的 F 检验

18.21 利用习题 18.19 与 18.20 的结果，针对表 18.10 的数据，在 $\alpha=0.05$ 水平下，对 $H_0:\mu_1=\mu_2=\mu_3$ 做右尾检验.

解 ☞ (1) $H_0:\mu_1=\mu_2=\mu_3$，H_1：三个均值不全相等；

(2) $\alpha=0.05$；

(3) 对于 $v_1=k-1=2$，$v_2=N-k=24$，利用附录表 A.8 可得临界值决策规则：若 $f^*>(f_{0.05,2,24}=3.4)$，则拒绝 H_0；

(4) 方差分析的结果见表 18.11. 由附录表 A.8 知 P 值范围为 $P<0.005$.

表 18.11

方差来源	平方和	自由度	均方	F
组间	0.291	2	0.146	7.68
组内	0.450	24	0.019	
合计	0.741	26		

(5) 因 $f^*>3.4$，故拒绝 H_0，从而接受 H_1. 此结论可由事实 $P<0.005$ 而确认. 从而，在 0.05 的显著水平下，有证据表明三个均值不全相等.

一般程序的方差分析：多重比较

18.22 我们在习题 18.21 中拒绝了 H_0，针对该数据(见表 18.10)，在同样的显著水平下，用 Kramer 推广的 Duncan 多重极差检验法确定诸均值的差异.

解 ☞ (1) 将 3 个样本均值从小到大排成一行

$$\begin{array}{ccc}\bar{x}_3 & \bar{x}_2 & \bar{x}_1\\ 1.09 & 1.23 & 1.35\end{array}$$

(2) 根据 $MSW=0.019$，又对于 $\alpha=0.05$ 及 $v=24$，在附录表 A.9 查得 r_p，则可由(18.49)式得到 R'_p 值，结果如下表

p	2	3
r_p	2.919	3.066
R'_p	0.402	0.423

(3) 当 $p=3$ 时，该子集合有最大，最小两个样本均值，由(18.50)式，其极差为

$$(\bar{x}_1-\bar{x}_3)\sqrt{\frac{2n_1n_3}{n_1+n_3}}=(1.35-1.09)\sqrt{\frac{144}{17}}=0.757$$

由于 $0.757 > R'_3 = 0.423$,故认为 $\bar{x}_1$ 与 $\bar{x}_3$ 显著不同,从而 $\mu_1 > \mu_3$.

(4) 由于在步骤(3)中获得了明显差异,则必须对该极差进行两两检验.

首先,

$$(\bar{x}_1 - \bar{x}_2)\sqrt{\frac{2n_1 n_2}{n_1 + n_2}} = (1.35 - 1.23)\sqrt{\frac{160}{18}} = 0.358$$

由于 $0.358 < R'_2 = 0.402$,故认为 $\bar{x}_1$ 与 $\bar{x}_2$ 无明显差异,从而 $\mu_1 = \mu_2$.

再者,

$$(\bar{x}_2 - \bar{x}_3)\sqrt{\frac{2n_2 n_3}{n_2 + n_3}} = (1.23 - 1.09)\sqrt{\frac{180}{19}} = 0.431$$

由于 $0.431 > R'_2 = 0.402$,故认为 $\bar{x}_2$ 与 $\bar{x}_3$ 差异显著,从而 $\mu_2 > \mu_3$.

(5) 综上所述,可知

$$\underline{\bar{x}_3} \quad \underline{\bar{x}_2 \quad \bar{x}_1}$$

因此可下结论:

$$\mu_3 < \mu_2 = \mu_1$$

一般程序的方差分析:置信区间的计算及方差齐性检验

18.23 在习题 18.22 中得结论 $\mu_3 < \mu_2 = \mu_1$,试根据数据计算 μ_3 的 90% 置信区间.

解 ☞ 由(18.26)式,有

$$\bar{x}_i \pm t_{\alpha/2,\nu}\sqrt{\frac{MSW}{n_i}}$$

由表 18.10 知 $\bar{x}_3 = 1.09$ 小时, $n_i = n_3 = 9$,由附录表 A.6 知 $t_{\alpha/2,\nu} = t_{0.10/2,24} = 1.711$,由习题 18.20 知 $MSW = 0.019$ 小时^2,则 μ_3 的 90% 置信区间为

$$1.09\text{ 小时} \pm 1.711\sqrt{\frac{0.019\text{ 小时}^2}{9}}$$

即

$$1.09\text{ 小时} \pm 0.079\text{ 小时}$$

18.24 针对表 18.10 的数据,用 Bartlett 法在 $\alpha = 0.05$ 水平下检验方差的齐性.

解 ☞ 根据上册的(7.16)式知,样本方差为 $s_1^2 = 0.0200$, $s_2^2 = 0.0201$, $s_3^2 = 0.0161$. 由(18.31)式, $k = 3$ 时,

$$N = \sum_{i=1}^{3} n_i = 8 + 10 + 9 = 27$$

由(18.30)式

$$s_p^2 = \frac{\sum_{i=1}^{3}(n_i - 1)s_i^2}{N - k} = \frac{0.4497}{24} = 0.0187$$

由(18.33)式

$$\begin{aligned} q &= (N - k)\log s_p^2 - \sum_{i=1}^{3}(n_i - 1)\log s_i^2 \\ &= [24 \times (-1.7282)] - (-41.5094) = 0.0326 \end{aligned}$$

由(18.34)式

$$h = 1 + \frac{1}{6}\left(\sum_{i=1}^{3}\frac{1}{n_i - 1} - \frac{1}{27 - 3}\right) = 1.0562$$

故由(18.32)式可得

$$b^* = 2.3026 \times \frac{0.0326}{1.0562} = 0.0711$$

方差齐性检验可归纳如下：

(1) $H_0:\sigma_1^2=\sigma_2^2=\sigma_3^2$，$H_1$：三个方差不全相等；

(2) $\alpha=0.05$；

(3) 在 $\alpha=0.05$，$v=k-1=3-1=2$ 时，查附表 A.7，利用 B 统计量可得临界值决策规则：若 $b^*>(\chi^2_{0.05,2}=5.99)$，则拒绝 H_0；

(4) B 统计量的样本观测值为 $b^*=0.0711$，且由附表 A.7 可见，当 $v=2$ 时，有一个介与 $\chi^2_{0.975}$ 与 $\chi^2_{0.950}$ 间的可供比较的 χ^{2*} 值，则 P 值范围为 $0.950<P<0.975$；

(5) 由于 $b^*<5.99$，故接受 H_0，即方差相同，此结论可由 $P>0.05$ 的事实来确认.

补 充 习 题

样本容量相等时的单向固定效应方差分析：H_0 与 H_1

18.25　某营养学家正在测试五种节食方案，标号为1,2,3,4,5. 他对25名妇女志愿者进行试验，将所有人随机地平分成5组，每组对应于一种方案. 在试验之初及按各方案进行两月之后测量了每位妇女的体重(精确到磅)，并计算每个人的重量差：起初的体重减掉后来的体重. 假定各重量差的总体服从正态分布，均值为 $\mu_1,\mu_2,\mu_3,\mu_4,\mu_5$，且各方差相等：$\sigma_1^2=\sigma_2^2=\sigma_3^2=\sigma_4^2=\sigma_5^2=\sigma^2$. 若他打算对这些数据做一个单向固定效应方差分析，则何为 H_0 与 H_1？

答案：$H_0:\mu_1=\mu_2=\mu_3=\mu_4=\mu_5$，$H_1$：5个均值不全相等

样本容量相等时的方差分析：整理数据

18.26　如下即是习题18.25的试验结果，它们表示了每位妇女的体重差：组1(6,7,6,8,7)，组2(8,9,10,8,9)，组3(7,9,8,6,9)，组4(14,13,12,15,14)，组5(12,13,11,11,14). 试由此数据，计算 $T_{1\cdot}$，$T_{2\cdot}$，$T_{3\cdot}$，$T_{4\cdot}$，$T_{5\cdot}$，$T_{\cdot\cdot}$.

答案：$T_{1\cdot}=34$，$T_{2\cdot}=44$，$T_{3\cdot}=39$，$T_{4\cdot}=68$，$T_{5\cdot}=61$，$T_{\cdot\cdot}=246$

18.27　由上题结果，求 $\bar{x}_{1\cdot}$，$\bar{x}_{2\cdot}$，$\bar{x}_{3\cdot}$，$\bar{x}_{4\cdot}$，$\bar{x}_{5\cdot}$，$\bar{x}_{\cdot\cdot}$.

答案：$\bar{x}_{1\cdot}=6.8$，$\bar{x}_{2\cdot}=8.8$，$\bar{x}_{3\cdot}=7.8$，$\bar{x}_{4\cdot}=13.6$，$\bar{x}_{5\cdot}=12.2$，$\bar{x}_{\cdot\cdot}=9.8$

样本容量相等时的方差分析：平方和

18.28　由18.26的结果，求 SST，SSA 与 SSW.

答案：$SST=195.4$，$SSA=171.0$，$SSW=24.4$

样本容量相等时的方差分析：均方

18.29　针对节食试验，根据上题结果确定 MSA 与 MSW.

答案：$MSA=42.75$，$MSW=1.22$

样本容量相等时的方差分析：F 检验

18.30　针对如上节食试验，利用上题结果，在 $\alpha=0.05$ 水平下，对 $H_0:\mu_1=\mu_2=\mu_3=\mu_4=\mu_5$ 做 F 检验. 给出 f^* 的值，说明接受还是拒绝 H_0，并给出当 H_0 为真时取到 f^* 值的近似概率.

答案：$f^*=35.04$；拒绝 H_0，因为 $f^*>2.87$，$P<0.005$

样本容量相等时的方差分析：多重比较

18.31　在上题中，我们拒绝了 $H_0:\mu_1=\mu_2=\mu_3=\mu_4=\mu_5$，在同样的显著水平 $\alpha=0.05$ 下，试对该组数据利用 Duncan 多重极差检验法确定各均值间的差异. 请给出 R_p 值，并叙述关于样本均值及总体均值的结论.

答案：$R_2=1.457$，$R_3=1.530$，$R_4=1.576$，$R_5=1.608$；$\underline{\bar{x}_1\ \bar{x}_3}\ \underline{\bar{x}_3\ \bar{x}_2}\quad \underline{\bar{x}_5\ \bar{x}_4}$，可以明确地作结论：$\mu_1<\mu_2<\mu_5=\mu_4$，但尚不能由此检验判定 μ_3 与其他均值的关系

样本容量相等时的方差分析：置信区间的计算及方差齐性检验

18.32　在上题中有结论 $\mu_1<\mu_2<\mu_5=\mu_4$. 试由数据计算 μ_4 与 μ_5 的总均值 μ_w 的95%置信区间.

答案：1.29 ± 0.73

18.33　根据习题18.26与18.27的结果，在 $\alpha=0.05$ 水平下，用 Bartlett 法检验方差的齐性. 请给出 s_1^2，s_2^2，s_3^2，s_4^2，s_5^2，s_p^2 及 b^* 的值. 说明接受还是拒绝 $H_0:\sigma_1^2=\sigma_2^2=\sigma_3^2=\sigma_4^2=\sigma_5^2$，并且指出当 H_0 为真时，取到 b^* 值的近似概率.

答案：$s_1^2=0.70, s_2^2=0.70, s_3^2=1.70, s_4^2=1.30, s_5^2=1.70, s_p^2=1.22, b^*=1.3962$，由于 $b^*<9.49$，故接受 H_0，$0.5<P<0.9$

一般程序的单向固定效应方差分析：H_0 与 H_1

18.34　假如在习题 18.25 与 18.26 所述的节食试验中，25 名妇女并未全部完成试验，而是缺了两个人的数据：组 1 中的 6 及组 4 的 15. 若正态假定及方差齐性假定保持不变，欲对这 23 个数据做一个单向固定效应方差分析，试问 H_0 及 H_1 如何表述？

答案：$H_0: \mu_1=\mu_2=\mu_3=\mu_4=\mu_5$，$H_1$：5 个均值不全相等

一般程序的方差分析：整理数据

18.35　相应于上题，其数据如下：组 1(7,6,8,7)；组 2(8,9,10,8,9)；组 3(7,9,8,6,9)；组 4(14,13,12,14)；组 5(12,13,11,11,14). 由此计算 $T_{1.}, T_{2.}, T_{3.}, T_{4.}, T_{5.}$ 与 $T_{..}$.

答案：$T_{1.}=28$, $T_{2.}=44$, $T_{3.}=39$, $T_{4.}=53$, $T_{5.}=61$, $T_{..}=225$

18.36　由上题结果，求 $\bar{x}_{1.}, \bar{x}_{2.}, \bar{x}_{3.}, \bar{x}_{4.}, \bar{x}_{5.}$ 与 $\bar{x}_{..}$.

答案：$\bar{x}_{1.}=7.0, \bar{x}_{2.}=8.8, \bar{x}_{3.}=7.8, \bar{x}_{4.}=13.2, \bar{x}_{5.}=12.2, \bar{x}_{..}=9.8$

一般程序的方差分析：平方和

18.37　针对习题 18.35 的结果计算 SST, SSA 与 SSW.

答案：$SST=153.9$, $SSA=132.8$, $SSW=21.1$

一般程序的方差分析：均方

18.38　由上题结果计算 MSA 与 MSW.

答案：$MSA=33.20, MSW=1.17$

一般程序的方差分析：F 检验

18.39　由上题结果，在 $\alpha=0.05$ 水平下，对 $H_0: \mu_1=\mu_2=\mu_3=\mu_4=\mu_5$ 做 F 检验. 给出 f^* 值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时，取 f^* 值的近似概率.

答案：$f^*<28.38$；拒绝 H_0，因为 $f^*>2.93$；$P<0.005$

一般程序的方差分析：多重比较

18.40　在习题 18.39 中，拒绝了 $H_0: \mu_1=\mu_2=\mu_3=\mu_4=\mu_5$，在同样显著水平之下，针对同一数据(见习题 18.35)利用 Kramer 的推广 Duncan 多重极差检验法确定各均值间的差异. 给出 R'_p 值，并说明关于样本均值及总体均值的结论.

答案：$R'_2=3.214, R'_3=3.373, R'_4=3.472, R'_5=3.541$，$\bar{x}_1\ \bar{x}_3\ \bar{x}_2\quad \bar{x}_5\ \bar{x}_4$，结论是：$\mu_1<\mu_2<\mu_5=\mu_4$，但尚不能由此检验确定 μ_3 与其他均值的关系.

一般程序的方差分析：置信区间的计算与方差齐性检验

18.41　在上题中有结论：$\mu_1<\mu_2<\mu_5=\mu_4$，由数据计算 $\mu_5-\mu_2$ 的 95% 置信区间.

答案：3.4 ± 1.44

18.42　根据习题 18.35 与 18.36 的结果，在 $\alpha=0.05$ 水平下，利用 Bartlett 法检验方差的齐性. 给出 $s_1^2, s_2^2, s_3^2, s_4^2, s_5^2, s_p^2$ 及 b^* 的值，说明接受还是拒绝 $H_0: \sigma_1^2=\sigma_2^2=\sigma_3^2=\sigma_4^2=\sigma_5^2$，并且给出在 H_0 为真时，能取到 b^* 值的近似概率.

答案：$s_1^2=0.67, s_2^2=0.70, s_3^2=1.70, s_4^2=0.92, s_5^2=1.70, s_p^2=1.18, b^*=1.4483$，因为 $b^*<9.49$，故接受 H_0，$0.5<P<0.9$

第十九章 回归和相关

19.1 两变量间关系的研究

到目前为止,我们所关注的均为**一元数据**,即样本观测值来自于一个变量.现在,我们要考虑**二元数据**,即样本观测值来自于两个变量.例如,对于甲虫总体,考虑体温和氧耗量,随机抽取容量为 n 的一个样本,对于每个个体,观测它的体温和氧耗量.

分析两个变量之间的关系,首先要构造样本数据的**散点图**.图 19-1 就是甲虫的体温和氧耗量的散点图.如果所有的点散布成一条直线,如图 19-1,我们说这两个变量有**线性关系**,否则,称这两个变量**不具有线性关系**.本章只研究线性关系.

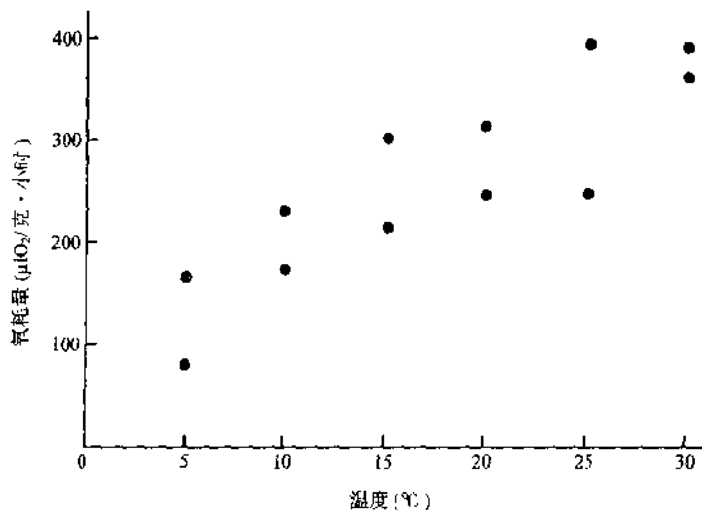


图 19-1

研究两个变量之间的关系,一般用**回归分析**或**相关分析**.在回归分析中,一个变量视为自变量,另外一个变量视为因变量(尽管它们之间可能不是因果关系,见上册的 1.19 节).自变量不一定是随机变量,可由研究者决定;因变量被假定为随机变量,通过抽样得到.于是,如果多次重复研究,自变量的取值可保持不变(完全由研究者设置),而因变量的值在抽样过程中将会发生变化.回归分析的目的就是要给出一个描述二者关系的数学方程.这个方程可以用来当已知自变量的值时,预测相应的因变量的值.回归分析的一个例子是:在试验中,养殖在实验室内的甲虫在不同环境温度(自变量)下的氧耗量(因变量)之间的关系,见图 19-1.而相关分析是研究两个自变量的相关程度,两个自变量均假定为随机变量.例如,在不同的环境条件下,野外的甲虫的氧耗量与体温的关系.由于两个变量的样本观测值均是通过抽样得到的,因而随着样本不同,观测值也不同.

19.2 简单线性回归模型

简单线性回归模型是分析自变量和因变量之间是否有直线型关系.术语"简单"是为区分双变量分析与多变量分析(见 19.11 节).习惯上,X 表示自变量,Y 表示因变量.若 X 与 Y 完全是直线型关系,则可表为

$$Y = a + bX \tag{19.1}$$

这里 a 为 Y 的截距，b 为斜率，它刻画了 X 与 Y 的关系. 常数 a，b 可由直线上的两点 (x_1, y_1)，(x_2, y_2) 确定. y 的截距 a 由 $X=0$ 时的 Y 值确定. 斜率 b 刻画了 Y 的改变量与 X 的相应改变量的比

$$b = \frac{y_2 - y_1}{x_2 - x_1} \tag{19.2}$$

当自变量 $X=x_i$ 已知时，(19.1)式可以用来预测因变量 Y 的相应值

$$y_i = a + bx_i \tag{19.3}$$

这里 a，b 为参数，常称为**回归系数**. 在实际研究中，Y 是不可能完全由 X 预测出的. 因为，对于任意给定的 x_i，对于 Y 的观测值，总存在随机误差. 我们所能做的是对每个 x_i，由样本观测值 y_i 估计 Y. 给定 x_i，Y 的期望值为

$$E(Y \mid x_i) = \mu_{Y \mid x_i} = a + bx_i \tag{19.4}$$

在线性回归中，我们用抽样数据估计 $\mu_{Y \mid x_i}$，a 和 b 的值，记总体均值 $\mu_{Y \mid x_i}$ 的估计为 $\hat{\mu}_{Y \mid x_i}$，样本均值为 $\hat{\mu}^*_{Y \mid x_i}$，a 的估计量记为 $\hat{a}$，其点估计值(样本值)记为 $\hat{a}^*$，b 的估计量记为 $\hat{b}$，其点估计值记为 $\hat{b}^*$，于是，由样本数据可得回归直线

$$\hat{\mu}^*_{Y \mid x_i} = \hat{a}^* + \hat{b}^* x_i \tag{19.5}$$

估计 a，b 的步骤如下：(1) 选取自变量 x 的 n 个值 $(x_1, x_2, \cdots, x_n)$，对于每个 x_i 取 Y 的一个样本；(2) 描出所有样本序对的散点图(每个 x_i 对应两个或多个 y)；(3) 若 X 与 Y 是线性关系，则确定 X 与 Y 的线性方程.

19.3　最小二乘回归直线

最小二乘回归直线(或**最佳拟合直线**)是一条直线，它使得所有点 (x_i, y_i) 相对于该直线的偏差的平方和尽可能达到最小，下面给出寻找这条直线的方法，即**最小二乘法**.

对于给定的 $X=x$，Y 的观测值不可能恰好等于由(19.1)式给出的预测值. 图 19-2 中的有箭头的线段则刻画了这种偏差. 所有 Y 的样本观测值的偏差的平方和(见 18.9 节)可以用来刻画通过这些数据点的直线"拟合"数据的程度.

我们根据 a，b 的点估计 $\hat{a}^*$ 与 $\hat{b}^*$ 来给出最小二乘回归直线的方程. (a，b 的点估计 $\hat{a}^*$ $\hat{b}^*$ 称为 a，b 的**最小二乘估计**)，不加推导，直接给出 a 的点估计

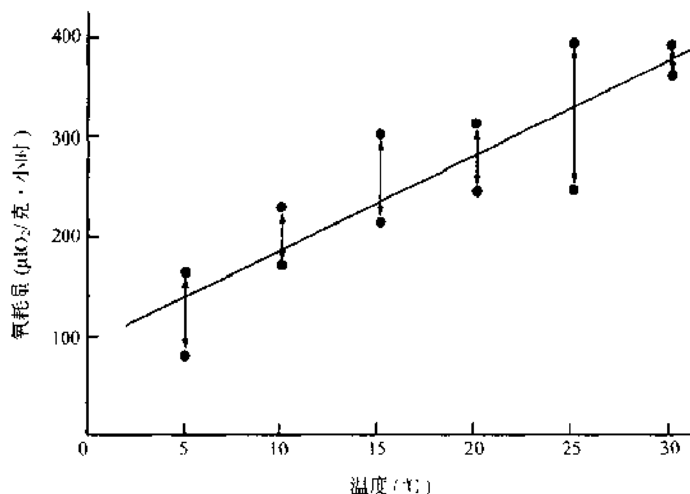


图 19-2

$$\hat{a}^* = \bar{y} - \hat{b}^* \bar{x} \tag{19.6}$$

和 b 的点估计

$$\hat{b}^* = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \tag{19.7}$$

这里 $\bar{y}$ 表示随机变量 Y 相对于每个 x 的样本观测值的平均值,$\bar{x}$ 表示 x 的平均值.由于这种方法得到的回归直线,仅对可观测的 x 适用,并非对所有的 x 都适用,因而不能外推到所用的 x 的值的范围之外.

点估计 $\hat{a}^*$,$\hat{b}^*$ 的计算公式均可利用下列表达式简化

$$\sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n} \tag{19.8}$$

$$\sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n} \tag{19.9}$$

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} \tag{19.10}$$

于是有

$$\hat{b}^* = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} \tag{19.11}$$

这里 n 表示序对的个数,再由(19.6)式可得 $\hat{a}^*$.

例 19.1 一位运动生理学家想根据训练水平(单位:千克-米/分钟,即每分钟将 1 千克的物体升高 1 米)来预测心脏血液输出量(每分钟由心脏输出的血液的升数).他选取四个训练水平:0,300,600,900(千克-米/分钟),随机抽取 20 人的一个样本,随机分成四组,每组 5 人,每个水平一组;训练 15 分钟后,测量他们的心脏血液输出量,结果如表 19.1 所示.求这两个变量之间的最小二乘回归直线?给定训练水平 700 千克-米/分钟,求心脏血液输出量的预测值?

表 19.1

个体	训练水平(千克-米/分钟)	心脏血液输出量(升/分钟)
1	0	4.4
2	0	5.6
3	0	5.2
4	0	5.4
5	0	4.4
6	300	9.1
7	300	8.6
8	300	8.5
9	300	9.3
10	300	9.0
11	600	12.8
12	600	13.4
13	600	13.2
14	600	12.6
15	600	13.2
16	900	17.0
17	900	17.3
18	900	16.5
19	900	16.8
20	900	17.2

解 首先利用(19.11)式计算点估计 $\hat{b}^*$. 为此,先计算 $\sum x_i$, $\sum x_i^2$, $\sum y_i$ 和 $\sum x_iy_i$(见表 19.2),又 $n=20$,于是有

$$\hat{b}^* = \frac{128,790 - \frac{(9,000)(219.5)}{20}}{6,300,000 - \frac{(9,000)^2}{20}} = 0.01334 \text{ 升 / 千克-米}$$

其次,由(19.6)式求得

$$\hat{a}^* = \frac{219.5}{20} - (0.01334)\frac{9,000}{20} = 4.972 \text{ 升 / 分钟}$$

(这里,Y 的截距 a 的估计值 $\hat{a}^*$ 四舍五入到比 Y 的观测值多一位小数,斜率 b 的估计值 $\hat{b}^*$ 四舍五入到当 x 的最大值多次计算时,比 Y 的观测值多一位小数.因而 $\hat{a}^*$ 为 4.97, $\hat{b}^*$ 为 0.0133).

最小二乘回归直线为: $\hat{\mu}^*_{Y|x} = 4.97 + 0.0133x$. 给定训练水平 700 千克-米/分钟,心脏血液输出量的预测值为 14.28 升/分钟.

表 19.2

	x_i	y_i	x_i^2	y_i^2	x_iy_i
	0	4.4	0	19.36	0
	0	5.6	0	31.36	0
	0	5.2	0	27.04	0
	0	5.4	0	29.16	0
	0	4.4	0	19.36	0
	300	9.1	90,000	82.81	2,730
	300	8.6	90,000	73.96	2,580
	300	8.5	90,000	72.25	2,550
	300	9.3	90,000	86.49	2,790
	300	9.0	90,000	81.00	2,700
	600	12.8	360,000	163.84	7,680
	600	13.4	360,000	179.56	8,040
	600	13.2	360,000	174.24	7,920
	600	12.6	360,000	158.76	7,560
	600	13.2	360,000	174.24	7,920
	900	17.0	810,000	289.00	15,300
	900	17.3	810,000	299.29	15,570
	900	16.5	810,000	272.25	14,850
	900	16.8	810,000	282.24	15,120
	900	17.2	810,000	295.84	5,480
$\sum$	9,000	219.5	6,300,000	2,812.05	128,790

19.4 方差 $\sigma^2_{Y\cdot X}$ 的估计

最小二乘回归直线的方差 $\sigma^2_{Y\cdot X}$,可采用如下方法估计:对于所有的 x_i, Y 相对于总体均值的偏差的平方和,作为其估计量.对于自变量 X 的每一个值,假定因变量 Y 的样本观测值服从均值为

$$\hat{\mu}_{Y|x} = \hat{a} + \hat{b}x \tag{19.12}$$

的正态分布,且对所有的 x_i, Y 的分布有相同的方差,该方差的一个无偏估计是

$$s^2_{y\cdot x} = \frac{\sum (y_i - \hat{\mu}_{Y|x_i})^2}{n-2} \tag{19.13}$$

即 y_i 相对于回归直线的均方差,这里脚标 $y\cdot x$ 表示该表达式是我们研究 y 对 x 的回归时 Y 的方差, $n-2$ 为样本容量 n 减去自由度 2,两个参数(a 和 b)的估计量的自由度各为 1.

(19.13)式可化为

$$s_{y\cdot x}^2=\frac{\sum y_i^2-\hat{a}^*\sum y_i-\hat{b}^*\sum x_iy_i}{n-2} \tag{19.14}$$

标准差为

$$s_{y\cdot x}=\sqrt{s_{y\cdot x}^2} \tag{19.15}$$

标准差(**估计量的标准误**),刻画了 y_i 相对于回归直线的偏差,它反映了由 X 来估计 Y 的不确定程度.

例 19.2　考虑例 19.1 中的试验,求方差的估计量和估计量的标准差?

解　将 $n=20$, $\hat{a}^*=4.97$ 升/分钟, $\hat{b}^*=0.0133$ 升/千克-米,及表 19.2 中的一些值代入(19.14)式,得方差的估计为

$$s_{y\cdot x}^2=\frac{2,812.05-(4.97)(219.5)-(0.0133)(128,790)}{20-2}=0.4571(\text{升}/\text{分钟})^2$$

由(19.15)式得该估计的标准差为

$$s_{y\cdot x}=\sqrt{0.4571}=0.6761\approx 0.68\ \text{升}/\text{分钟}$$

(我们将标准差,或估计的标准误四舍五入到比 Y 的观测值多一位小数).

19.5　截距 $\hat{a}$ 和斜率 $\hat{b}$ 的均值和方差

假定所有的 x_i, y_i 均服从正态分布(见 19.4 节),由于估计量 $\hat{a}$, $\hat{b}$ 为这些 Y_i 值的线性函数,故 $\hat{a}$, $\hat{b}$ 亦服从正态分布,从而可求得 $\hat{a}$, $\hat{b}$ 的方差.

不加推导,直接给出结论:估计量 $\hat{a}$ 的均值为 $\mu_{\hat{a}}=a$,方差为

$$s_{\hat{a}}^2=s_{y\cdot x}^2\left(\frac{1}{n}+\frac{\bar{x}^2}{\sum(x_i-\bar{x})^2}\right) \tag{19.16}$$

这里 $s_{y\cdot x}^2$ 是总体方差 $\sigma_{Y\cdot X}^2$ 的最佳估计[见(19.13),(19.14)式].注意到(19.8)式,则(19.16)式可化为

$$s_{\hat{a}}^2=s_{y\cdot x}^2\left[\frac{1}{n}+\frac{\left(\frac{\sum x}{n}\right)^2}{\sum x^2-\frac{(\sum x)^2}{n}}\right] \tag{19.17}$$

不加推导,直接给出结论:估计量 $\hat{b}$ 的均值为 $\mu_{\hat{b}}=b$,方差为

$$s_{\hat{b}}^2=\frac{s_{y\cdot x}^2}{\sum(x_i-\bar{x})^2} \tag{19.18}$$

这里 $s_{y\cdot x}^2$ 为总体方差 $\sigma_{Y\cdot X}^2$ 的最佳估计[见(19.13),(19.14)式].(19.18)式可化为

$$s_{\hat{b}}^2=\frac{s_{y\cdot x}^2}{\sum x^2-\frac{(\sum x)^2}{n}} \tag{19.19}$$

这些估计量的标准差为

$$s_{\hat{a}}=\sqrt{s_{\hat{a}}^2} \tag{19.20}$$

及

$$s_{\hat{b}} = \sqrt{s_{\hat{b}}^2} \tag{19.21}$$

例 19.3　考虑例 19.1 中的试验，求 $\hat{a}$，$\hat{b}$ 点估计的标准差？

解　将 $s_{y,x}^2 = 0.4571$，$n=20$，及表 19.2 中的一些值，代入(19.17)和(19.20)式，得 $\hat{a}^*$ 的标准差

$$s_{\hat{a}^*} = \sqrt{s_{\hat{a}^*}^2} = \sqrt{0.4571\left[\frac{1}{20} + \frac{\left(\frac{9,000}{20}\right)^2}{6,300,000 - \frac{(9,000)^2}{20}}\right]} = \sqrt{0.063994} = 0.25297 \approx 0.253 \text{ 升 / 分钟}$$

类似地，由(19.18)和(19.21)式可得 $\hat{b}^*$ 的标准差

$$s_{\hat{b}^*} = \sqrt{s_{\hat{b}^*}^2} = \sqrt{\frac{0.4571}{6,300,000 - \frac{(9,000)^2}{20}}} = \sqrt{0.0000002032} = 0.0004507 \approx 0.00045 \text{ 升 / 千克-米}$$

19.6　截距 a 和斜率 b 的置信区间

建立最小二乘回归直线后，接下来，我们想寻找一个范围，使得我们有很大把握断定参数真值落入其中．由 19.5 节，我们知道 y 的截距 a 的估计量 $\hat{a}$ 和斜率 b 的估计量 $\hat{b}$ 可以假定服从均值分别为 a，b 的正态分布．在此假定下，可利用 t 分布，求出 a，b 的置信区间．

不加推导，我们说，统计量

$$T = \frac{\hat{a} - a}{S_{\hat{a}}} \tag{19.22}$$

的抽样分布是自由度为 $v = n-2$ 的 t 分布，(这里 $S_{\hat{a}}$ 是一个变量，可取值 $s_{\hat{a}}$)这个统计量可假定取特定值

$$t = \frac{\hat{a}^* - a}{s_{\hat{a}^*}}$$

对于这个 t 统计量，我们知道[见(14.15)式]

$$P\left(-t_{\alpha/2,v} \leqslant \frac{\hat{a} - a}{S_{\hat{a}}} \leqslant t_{\alpha/2,v}\right) = 1 - \alpha$$

等价地(见 14.8 节)有

$$P(\hat{a} - t_{\alpha/2,v}S_{\hat{a}} < a < \hat{a} + t_{\alpha/2,v}S_{\hat{a}}) = 1 - \alpha \tag{19.23}$$

从而，y 的截距 a 的 $(1-\alpha)100\%$ 的置信限为

$$[\hat{a}^* - t_{\alpha/2,v}s_{\hat{a}^*}, \hat{a}^* + t_{\alpha/2,v}s_{\hat{a}^*}]$$

类似地，可导出斜率 b 的一个置信区间．不加推导，我们说，统计量

$$T = \frac{\hat{b} - b}{S_{\hat{b}}} \tag{19.24}$$

的抽样分布是自由度为 $v = n-1$ 的 t 分布(这里 $S_{\hat{b}}$ 是一个变量，可以取值 $s_{\hat{b}}$)．对于此 t 统计量，有[见(14.15)式]

$$P\left(-t_{\alpha/2,v} \leqslant \frac{\hat{b} - b}{S_{\hat{b}}} \leqslant t_{\alpha/2,v}\right) = 1 - \alpha$$

或

$$P(\hat{b} - t_{\alpha/2,v}S_{\hat{b}} < b < \hat{b} + t_{\alpha/2,v}S_{\hat{b}}) = 1 - \alpha \tag{19.25}$$

于是，斜率 b 的 $(1-\alpha)100\%$ 置信区间为

$$\left[\hat{b}^* - t_{\alpha/2,v}s_{\hat{b}^*},\ \hat{b}^* + t_{\alpha/2,v}s_{\hat{b}^*}\right]$$

对于给定样本，在给定置信水平和自由度的前提下，由 t 分布表(附录表 A.6)可得临界值，从而得到所需的置信区间.

例 19.4　考虑例 19.1 中的试验，求 a，b 的 95%置信区间？

解　由例 19.1 知 $\hat{a}^*=4.97$ 升/分钟，且由例 19.3 知 $s_{\hat{a}^*}=0.253$ 升/分钟. 再由附录表 A.6，对于 $v=n-2=18$，有 $t_{0.05/2,18}=2.101$. 从而利用(19.23)式，可得 a 的 95%置信区间为

$$4.97\ 升/分钟 \pm (2.101\times 0.253\ 升/分钟)$$

$$4.97\ 升/分钟 \pm 0.532\ 升/分钟$$

类似地，由例 19.1 知 $\hat{b}^*=0.0133$ 升/千克-米，且由例 19.3 知 $s_{\hat{b}^*}=0.00045$ 升/千克-米. 由上面的自由度 $v=18$ 及相同的临界值，利用(19.25)式，可得 b 的 95%置信区间为

$$(0.0133\ 升/千克\text{-}米)\pm(2.101\times 0.00045\ 升/千克\text{-}米)$$

$$(0.0133\ 升/千克\text{-}米)\pm(0.00095\ 升/千克\text{-}米)$$

19.7　方差 $\sigma_{Y\cdot X}^2$ 的置信区间

给定 x 和 y 的样本序对，Y 分布的方差可由 $s_{y\cdot x}^2$ 给出[见(19.13)，(19.14)式]，于是该样本方差即为总体方差 $\sigma_{Y\cdot X}^2$ 的点估计.

不加推导，我们说，统计量

$$X^2=\frac{(n-2)S_{Y\cdot X}^2}{\sigma_{Y\cdot X}^2} \tag{19.26}$$

的抽样分布是自由度为 $v=n-2$ 的 χ^2 分布(这里 $S_{Y\cdot X}$ 是一个变量，可取值 $s_{y\cdot x}$). 由 χ^2 分布知[见(15.6)式]

$$P\left[\chi_{1-\alpha/2,v}^2\leqslant\frac{(n-2)S_{Y\cdot X}^2}{\sigma_{Y\cdot X}^2}\leqslant\chi_{\alpha/2,v}^2\right]=1-\alpha$$

等价地(见 15.5 节)，有

$$P\left[\frac{(n-2)S_{Y\cdot X}^2}{\chi_{\alpha/2}^2}\leqslant\sigma_{Y\cdot X}^2\leqslant\frac{(n-2)S_{Y\cdot X}^2}{\chi_{1-\alpha/2,v}^2}\right]=1-\alpha \tag{19.27}$$

因而 $\sigma_{Y\cdot X}^2$ 的置信区间为

$$\left[\frac{(n-2)s_{y\cdot x}^2}{\chi_{\alpha/2,v}^2},\frac{(n-2)s_{y\cdot x}^2}{\chi_{1-\alpha/2,v}^2}\right]$$

对于给定样本，在已知自由度 $v=n-2$ 和置信水平 α 的前提下，查 χ^2 分布表(表 A.7)可得临界值，从而求得置信区间.

总体标准差 $\sigma_{Y\cdot X}$ 的置信区间可通过将 $\sigma_{Y\cdot X}^2$ 的置信区间的两个端点值直接开方而得.

例 19.5　考虑例 19.1 中的试验，求训练水平与心脏血液输出量之间线性关系的总体方差 $\sigma_{Y\cdot X}^2$ 的 95%置信区间？

解　已知 $n=20$，由例 19.2 得 $s_{y\cdot x}^2=0.4571$(升/分钟)^2. 此外 $\alpha/2=0.025$，$1-\alpha/2=0.975$，由表 A.7 得：$\chi_{0.025,18}^2=31.35$，$\chi_{0.975,18}^2=8.23$. 再由(19.27)式，求得 $\sigma_{Y\cdot X}^2$ 的 95%置信区间为

$$\left[\frac{(18\times 0.4571)}{31.53},\frac{(18\times 0.4571)}{8.23}\right]$$

或

$$[0.261(升/分钟)^2,1.000(升/分钟)^2]$$

19.8 Y 期望值的预测区间

给定 X，Y 的期望值可由回归直线

$$E(Y \mid x = x_0) = \mu_{Y|x_0} = a + bx_0 \tag{19.28}$$

得到，这里 x_0 为 X 的给定值，根据 x_0 来预测 Y 的值. 所谓**预测区间**，即为寻找一个范围，使得我们可以以 $(1-\alpha)100\%$ 的置信度来预测 Y 的均值，该区间可由 Y 的均值的标准差（见13.11 节）（通常记为 $s_{\hat{\mu}_{Y|x_0}}$）而得，即

$$s_{\hat{\mu}_{Y|x_0}} = s_{y\cdot x}\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2}} \tag{19.29}$$

这里 $s_{y\cdot x}$ 是 y_i 相对于回归直线的标准差[见(19.15)式]，x_0 是 X 的给定值，由此来预测 Y 的值.

对于小样本($n<30$)，由自由度为 $v=n-2$ 的 t 分布可得预测区间

$$\hat{\mu}_{Y|x_0} \pm t_{\alpha/2,v}s_{\hat{\mu}_{Y|x_0}}$$

对于大样本($n \geqslant 30$)，由 Z 分布，可得预测区间. 此外，对于大样本，关于 $s_{\hat{\mu}_{Y|x_0}}$ 的表达式可简化为

$$s_{\hat{\mu}_{Y|x_0}} = \frac{s_{y\cdot x}}{\sqrt{n}}$$

预测区间为

$$\hat{\mu}_{Y|x_0} \pm z_{\alpha/2}\left(\frac{s_{y\cdot x}}{\sqrt{n}}\right)$$

例 19.6 考虑例 19.1 中的试验，试利用所确定的最小二乘回归直线，预测训练水平为 750 千克-米/分钟时的心脏血液输出量及其 95% 预测区间.

解 由例 19.1 知 $\hat{\mu}_{Y|x} = 4.97 + 0.0133x$. 给定训练水平 $x = 750$ 千克-米/分钟，由此方程得心脏血液输出量(Y)的预测值为

$$\hat{\mu}_{Y|x=750} = 4.97 + 0.0133(750) = 14.945 \approx 14.95 \text{ 升 / 分钟}$$

由于 $n<30$，故采用 t 分布求此值的预测区间. 首先利用(19.29)式计算 Y 的期望值的标准差. 由例 19.2 得 $s_{y\cdot x} = 0.68$ 升/分钟，利用(19.8)式及表 19.2 中的一些值，可得

$$s_{\hat{\mu}_{Y|x_0}} = 0.68\sqrt{1 + \frac{1}{20} + \frac{\left(750 - \frac{9,000}{20}\right)^2}{6,300,000 - \frac{(9,000)^2}{20}}} = 0.7099 \approx 0.71 \text{ 升 / 分钟}$$

对于 $\hat{\mu}_{Y|x=750} = 14.95$ 升/分钟，$s_{\hat{\mu}_{Y\cdot x_0}} = 0.71$ 升/分钟，及 $t_{0.05/2,18} = 2.101$（查表 A.6 得到），可得预测区间为

$$14.95 \text{ 升 / 分钟} \pm (2.101 \times 0.71 \text{ 升 / 分钟})$$
$$14.95 \text{ 升 / 分钟} \pm 1.492 \text{ 升 / 分钟}$$

19.9 关于斜率 b 的假设检验

在采用回归分析研究两个变量之间的关系时，我们假定其中一个变量(Y)为因变量，而另外一个变量(X)为自变量，这个假设可通过观察回归直线的斜率是否显著为零来检验，如果假设错误，即 Y 与 X 无关，则 Y 的分布变化不与 X 的变化一致. 因此，对于每个 X 的值，Y 的均值不变，从而斜率 b 为零.

对于给定 x,由回归直线 $\hat{\mu}_{Y|x}=\hat{a}+\hat{b}x$ 可得 Y 的均值,若 $b\neq0$,则对于每个 x,Y 的期望值都将不同;反之,若 $b=0$,则对于每个 x,Y 的期望值都为 a,于是我们可检验原假设 $H_0:b=0$,和下列对立假设之一:$H_1:b\neq0$,$H_1:b>0$,或 $H_1:b<0$.

已知 $\hat{b}$ 服从均值为 b 的正态分布(见 19.5 节),因此,当 $b=0$ 时,由于 t 统计量[见(19.24)式]服从自由度为 $v=n-2$ 的 t 分布,对应于备择假设的决策规则为

对于 $H_1:b\neq0$,若 $t^*<-t_{\alpha/2,v}$ 或 $t^*>t_{\alpha/2,v}$,则拒绝 H_0;

对于 $H_1:b>0$,若 $t^*>t_{\alpha,v}$,则拒绝 H_0;

对于 $H_1:b<0$,若 $t^*<-t_{\alpha,v}$,则拒绝 H_0.

例 19.7　由例 19.1 知最小二乘回归直线的斜率为 $\hat{b}^*=0.0133$ 分钟/千克-米;由例 19.3 知 $\hat{b}$ 的标准差为 0.00045 分钟/千克-米. 给定 $\alpha=0.05$,采用临界值决策规则,对零假设 $H_0:b=0$,做右侧检验.

解 ☞　(1) $H_0:b=0$,　$H_1:b>0$;

(2) $\alpha=0.05$;

(3) 利用(19.24)式定义的检验统计量,给定 $\alpha=0.05$,查附录表 A.6,对于右侧检验,临界值为 $t_{0.05,18}=1.734$. 因而,决策规则为:若 $t^*>1.734$,则拒绝 H_0;

(4) t 的样本值为

$$t^*=\frac{0.0133-0}{0.00045}=29.556$$

查附录表 A.6 知,我们得到 $t^*=29.556$ 的近似概率为 $P<0.0005$;

(5) 由于 $t^*>1.734$,故拒绝 H_0,接受 H_1. 这一结论可由下述事实证实:得到 $t^*=29.556$ 的概率小于 0.05,因而心脏血液输出量依赖于训练水平.

19.10　两样本或多样本的简单线性回归方程的比较

对两样本或多样本,已经得到多个线性回归方程,检验这些回归方程是否显著不同的统计方法已有许多,一些方法是检验斜率的差异,而另外一些方法是检验 y 的截距的差异. 当只比较两个回归方程时,检验统计量是建立在 t 分布基础上的;而当比较多个回归方程时,检验统计量则基于 F 分布基础之上. 对于每一种情形,我们的目的都是判断这些样本是否来自于同一个总体. 关于这些方法的深入讨论,已超出本书范围,故略去.

19.11　多元线性回归

对于很多统计问题而言,我们所感兴趣的变量是由多个因子共同影响的,比如,人的身高是由基因和营养共同影响的. 一个因变量与两个或多个自变量之间的线性关系,可利用**多元线性回归**来分析. 正如简单线性回归一样,我们假定因变量是随机变量,且样本数据来自于对每个自变量都有相同方差的正态总体.

刻画一个因变量与两个或多个自变量之间的关系的多元线性回归方程是一个多项式方程(见上册 1.11 节). 这些多元方程有最简形式,即对于因变量 Y 和两个自变量 x,x',回归方程为

$$\mu_{Y|x,x'}=a+bx+b'x'$$

正如简单线性回归一样,a,b,b' 的值可通过使平方和

$$\sum(y_i-\hat{a}^*-\hat{b}^*x_i-\hat{b}'^*x_i')^2$$

尽可能达到最小而求得,即可由样本观测值求得最小二乘回归直线.

由 Y 的样本观测值,亦可求得这些参数的最佳估计,但就多元回归分析而言,一般计算过

程较复杂,可由计算机操作完成.由于方程的复杂性及需要用计算机分析,本书不对多元回归分析更多涉猎.

19.12 简单线性相关

对于一个容量为 n 的样本观测对(x_i, y_i),简单线性回归与简单线性相关是分析变量 X 与 Y 之间关系的两种不同方法.对于简单线性回归[正如 19.2 节至 19.10 节所描述的],变量 X 是固定的(由试验者设定),而 Y 是随机变量,线性回归的目的就是建立一个线性方程,当已知 X 的值时,可由该方程预测出 Y 的期望值;而对于简单线性相关,X 与 Y 均为随机变量,我们的目的是确定它们之间线性相关的程度.

两个随机变量之间的关系可由散点图(图 19-3)看出.散点图上的每个点都由一对观测值确定,一个是 X 的值,另一个是 Y 的值,两个变量的相关程度也就是各数据点聚集在一条直线周围的程度.正相关[图 19-3(a)]表明两个变量之间有着同向关系:一个变量增大时,另一个变量也增大;负相关[图 19-3(b)]表明两个变量之间有反向关系:一个变量增大时,另一个变量减小,当所有点都在一条直线上时,两个变量之间有完全线性关系(或者同向,或者反向).不相关表明二者之间没有关系[图 19-3(c)],或者有关系但不是线性的.

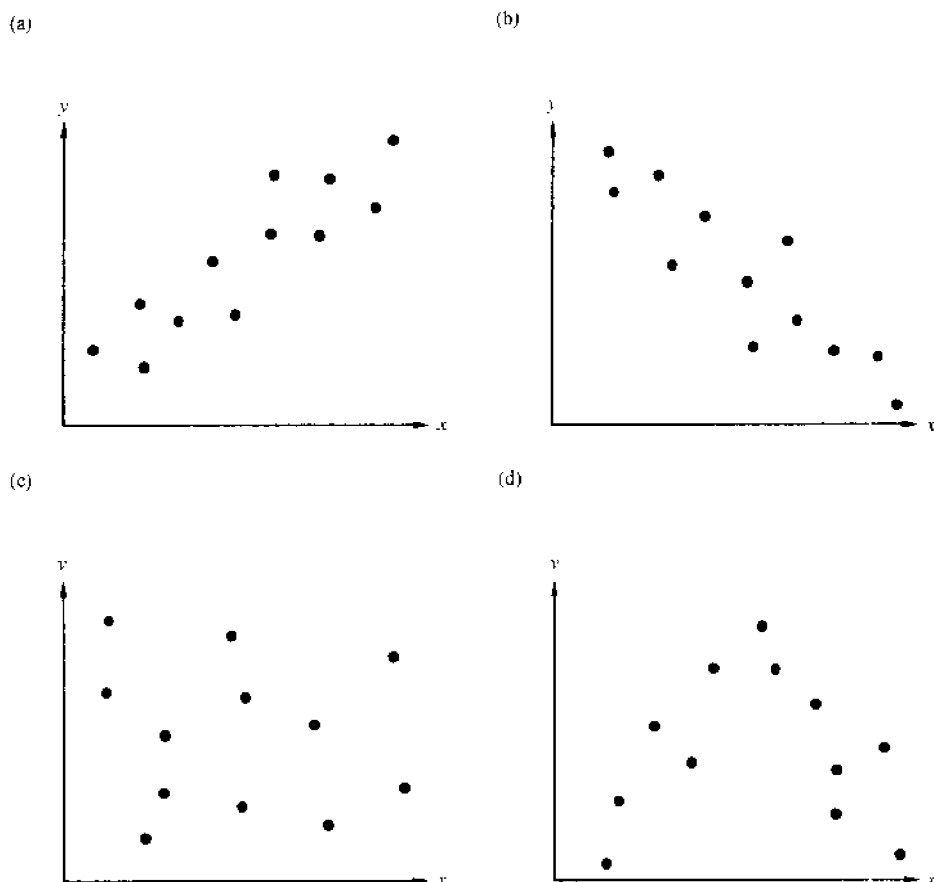


图 19-3

本章以后所讲的简单线性相关,均假定随机变量 X, Y 具有**联合概率分布** $f(x, y)$[见 11.13 节],一般假定为**二元正态分布**.这个假定意味着当 X, Y 为一对随机变量时,其联合分布具有光滑对称的山峰形状(确切形状依赖于相关程度).当 X 或 Y 的分布严重偏斜时,正态性的假定才不成立.

19.13 相关系数 r 的导出

在相关分析中,我们估计了总体参数 ρ,它刻画总体中的变量 X,Y 间的相关程度. ρ 的最常用的点估计是 Pearson **乘积矩相关系数或简单线性相关系数样本相关系数**,常记为 r,由于它是由英国统计学家 Karl Pearson(1857—1936)年给出的,故以 Pearson 命名. 为得到该统计量的表达式,我们需先介绍协方差的概念.

协方差刻画了两个随机变量相对于它们均值的同时偏差,它反映了两个变量共同变化的程度. 假定我们感兴趣的是随机变量 X 与 Y 的相关程度,并得到一些样本点;如图 19-4(a)或(b). 图中的每个点(x_i,y_i)均可根据它们相对于样本均值 $\bar{x}$ 和 $\bar{y}$ 的偏差来描述. 在本图中表为垂直点列和水平点列. 对每个样本点,两个偏差 $x_i-\bar{x}$ 与 $y_i-\bar{y}$ 相乘,然后将所有的乘积求做和,即得协方差. 为消除样本容量的影响,可将这些乘积的和被 $n-1$[n 为样本数据对(x_i,y_i)的个数]相除,即样本协方差为

$$\mathrm{cov}_{xy}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n-1} \tag{19.30}$$

协方差不能直接用来度量两个变量的相关程度,因为它的值与测量单位有关. 当两个变量的测量单位不同时(比如:身高的单位是英寸,而体重的单位是磅)会带来一些问题. 此外,测量单位还决定了协方差的数量级,比如,假定树的高度和直径是以厘米为单位测量的,它们的协方差将会大于以米为测量单位时的协方差. 因此,需将协方差标准化以消除测量单位的影响. Pearson 乘积矩相关,就是将协方差除以两个标准差 s_x,s_y 的乘积而进行标准化的

$$r=\frac{\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n-1}}{\sqrt{\dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}\sqrt{\dfrac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}}}$$

或

$$r=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2}\sqrt{\sum(y_i-\bar{y})^2}} \tag{19.31}$$

这样,测量单位被消去了,没有了单位的影响,r 仅刻画了两个变量之间的相关程度.

相关系数 r 的取值范围为$[-1,1]$,相关系数取正值或负值完全取决于分子. 考虑分子,当 x,y 的偏差按同方向变化,即 x 大于均值 $\bar{x}$ 时,y 亦大于其均值 $\bar{y}$ [见图 19-4(a)],则乘积$(x_i-\bar{x})(y_i-\bar{y})$为正,因为两个正数乘积为正,两个负数乘积亦为正;如果两个偏差按相反方向变化,即当 x 大于均值 $\bar{x}$ 时,y 小于其均值 $\bar{y}$ [见图 19-4(b)],则乘积$(x_i-\bar{x})(y_i-y)$为负;如果偏差$(x_i-\bar{x})$与偏差$(y_i-\bar{y})$的变化并不始终保持一致,则由于正、负号的偶然抵消,从而导致 $r=0$.

当所有的样本点都在一条直线上,且斜率为正时,r 取最大值 $+1$,这表明在(19.31)式中$(y_i-\bar{y})$可以用$(x_i-\bar{x})$代替;从而由(19.31)式得

$$r=\frac{\sum(x_i-\bar{x})(x_i-\bar{x})}{\sqrt{\sum(x_i-\bar{x})^2}\sqrt{\sum(x_i-\bar{x})^2}}=\frac{\sum(x_i-\bar{x})^2}{\sum(x_i-\bar{x})^2}=1$$

当所有的样本点都在一条直线上,且斜率为负时,r 取最小值 -1,这表明在(19.31)式中

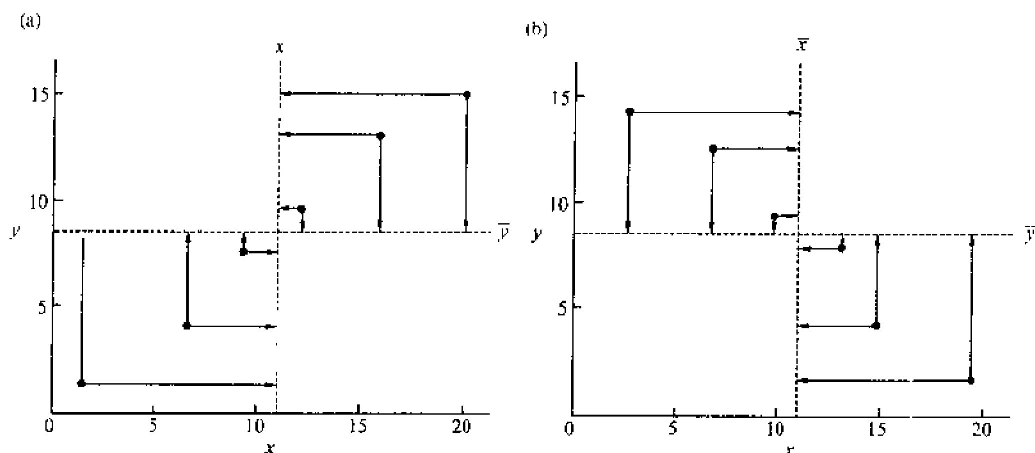


图 19-4

$(y_i-\bar{y})$可以用$-(x_i-\bar{x})$代替,从而由(19.31)式得 $r=-1$.

将(19.8),(19.9)及(19.10)式代入(19.37)式,可得

$$r=\frac{\sum xy-\frac{\sum x\sum y}{n}}{\sqrt{\left(\sum x^2\right)-\frac{\left(\sum x\right)^2}{n}}\sqrt{\left(\sum y^2\right)-\frac{\left(\sum y\right)^2}{n}}} \tag{19.32}$$

例 19.8　已知大黄蜂在飞行时翅膀肌肉的温度将会升高,一位昆虫学家想知道翅膀肌肉的温度是否与肌肉的工作量之间有线性关系.他用大黄蜂胸部(翅膀肌肉所在位置)的温度作为翅膀肌肉温度的一个标志,用腹部的重量作为飞行时肌肉工作量的一个标志,随机抽取 20 只大黄蜂作为一个样本,在飞行后,测量每只大黄蜂的胸部温度(单位:℃)和腹部重量(单位:毫克),测量结果见表 19.3.试估计这两个变量之间的线性相关系数?

表 19.3

黄蜂	腹部重量(毫克)	胸部温度(℃)
1	101.6	37.0
2	240.4	39.7
3	180.9	40.5
4	390.2	42.6
5	360.3	42.0
6	120.8	39.1
7	180.5	40.2
8	330.7	37.8
9	395.4	43.1
10	194.1	40.2
11	135.2	38.8
12	210.0	41.9
13	240.6	39.0
14	145.7	39.0
15	168.3	38.1
16	192.8	40.2
17	305.2	43.1
18	378.0	39.9
19	165.9	39.6
20	303.1	40.8

解 由于可假定大黄蜂总体的体温和重量服从二元正态分布,故可用样本相关系数 r^* 估计总体相关系数 ρ. 为此需先计算 $\sum x, \sum x^2, \sum y, \sum y^2, \sum xy$, 这些值见表 19.4. 将这些值代入(19.32)式得

$$r^* = \frac{192,154.74 - \left[\frac{(4,739.7)(802.6)}{20}\right]}{\sqrt{(1,295,879.09) - \frac{(4,739.7)^2}{20}}\sqrt{(32,264.56) - \frac{(802.6)^2}{20}}} = 0.626$$

样本相关系数 $r^* = 0.626$ 表明胸部温度与腹部重量之间有正向的关系.

表 19.4

	x_i	y_i	x_i^2	y_i^2	x_iy_i
	101.6	37.0	10,322.56	1,369.00	3,759.20
	240.4	39.7	57,792.16	1,576.09	9,543.88
	180.9	40.5	32,724.81	1,640.25	7,326.45
	390.2	42.6	152,256.04	1,814.76	16,622.52
	360.3	42.0	129,816.09	1,764.00	15,132.60
	120.8	39.1	14,592.64	1,528.81	4,723.28
	180.5	40.2	32,580.25	1,616.04	7,256.10
	330.7	37.8	109,362.49	1,428.84	12,500.46
	395.4	43.1	156,341.16	1,857.61	17,041.74
	194.1	40.2	37,674.81	1,616.04	7,802.82
	135.2	38.8	18,279.04	1,505.44	5,245.76
	210.0	41.9	44,100.00	1,755.61	8,799.00
	240.6	39.0	57,888.36	1,521.00	9,383.40
	145.7	39.0	21,228.49	1,521.00	5,682.30
	168.3	38.1	28,324.89	1,451.61	6,412.23
	192.8	40.2	37,171.84	1,616.04	7,750.56
	305.2	43.1	93,147.04	1,857.61	13,154.12
	378.0	39.9	142,884.00	1,592.01	15,082.20
	165.9	39.6	27,522.81	1,568.16	6,569.64
	303.1	40.8	91,869.61	1,664.64	12,366.47
$\sum$	4,739.7	802.6	1,295,879.09	32,264.56	192,154.74

19.14　总体相关系数 ρ 的置信区间

为确定总体参数 ρ 的 $(1-\alpha)100\%$ 置信区间, R. A. Fisher(见 17.17 节)给出了一种方法, 它涉及到一个新的变量 Z_r. 首先,利用下述变换可将样本相关系数 r 转化成相应的 z_r

$$z_r = 0.5\log_e\left(\frac{1+r}{1-r}\right)$$

这里 $\log_e$ 表示自然对数(见上册的习题 1.25). 幸运的是,对于这种变换,我们有表可以使用,例如附录中的表 A.10(**r 到 z_r 的变换**),可利用此表进行转换. 比如,已知 $r=0.634$,求相应的 z_r. 首先从表 A.10 的左首第一列找 0.63,然后从上方第一行找 0.004,行与列交叉处即为相应的 z_r 值: 0.7481. 表 A.10 只给出了正的 r 值,然而由于 r 及 z_r 的分布均关于原点对称,因而由表 A.10 也可查到 r 取负值时相应的 z_r 值.

通过这种变换,r 的抽样分布转化成 z_r 的抽样分布,其标准误为

$$\sigma_{z_r} = \sqrt{\frac{1}{n-3}} \tag{19.33}$$

这里 n 为样本对个数;其次,若假定随机变量 Z_r 近似服从正态分布,则下述统计量近似服从标准正态分布

$$Z = \frac{Z_r - z_\rho}{\sigma_{z_r}} \tag{19.34}$$

这里 z_ρ 为 ρ 的相应的 z_r 变换值;最后,将 Z 统计量代入(14.4)式,有

$$P(-z_{\alpha/2} \leqslant Z \leqslant z_{\alpha/2}) = 1 - \alpha$$

再进行通常的数学推导(见 14.8 节),可得 z_ρ 的近似置信限为

$$\left(z_r - z_{\alpha/2}\sqrt{\frac{1}{n-3}},\ z_r + z_{\alpha/2}\sqrt{\frac{1}{n-3}}\right)$$

一旦关于 z_r 的置信限得到,我们可利用表 A.10 将它转化成关于 r 值的置信限.正如所见到的那样,在解决问题时,z_r 的置信限是对称的,而 ρ 的置信区间(在将 z_r 转化成 r 之后),一般不是对称的.此外,置信区间长度(估计精度的标志)依赖于标准误 σ_r,而这又受到样本容量 n 的影响,样本越大,标准误越小,置信限也就越窄.

例 19.9　利用例 19.8 中得到的样本相关系数,求总体相关系数的近似 95%置信区间?

解　首先,将相关系数 r 转化成相应的 z_r 值.由表 A.10,对于 $r^*=0.626$,相应的 $z_r=0.7348$,置信区间为$[0.7348-z_{\alpha/2}\sigma_{z_r},\quad 0.7348+z_{\alpha/2}\sigma_{z_r}]$.利用 $z_{0.05/2}=1.96$, $n=20$,及

$$\sigma_{z_r} = \sqrt{\frac{1}{n-3}} = \sqrt{\frac{1}{17}} = 0.2425$$

得近似置信限为

$$[0.7348-(1.96\times0.2425),\quad 0.7348+(1.96\times0.2425)]$$

$$(0.2595,\quad 1.2101)$$

再由表 A.10,将这些置信限转化成相应的 r 值,于是当 $r=0.626$ 时的 95%置信区间为

$$(0.254,\quad 0.837)$$

19.15　用 r 分布检验总体相关系数 ρ 的假设

当已求得 r 的值时,一个重要问题是需要知道这个 r 值是真正刻画了两个变量 X,Y 之间具有的线性相关,还是仅仅由于样本的偶然相关而引起的.如果这两个变量并不是线性相关,则总体相关系数 ρ 为零,于是,可检验零假设 $H_0:\rho=0$ 和下列备择假设之一:$H_1:\rho\neq0$, $H_1:\rho>0$,或 $H_1:\rho<0$.该检验的逻辑原则,同其他检验相同.其目的是确定不同于零假设给出的 θ_0 的可由样本得到的点估计 $\hat{\theta}^*$ 的概率.对于相关系数,这一陈述为

$$P\{\text{不同于 0 的 }\rho\text{ 的点估计 }r^* \mid H_0\text{ 为真}\}$$

回忆 16.3 节,一个检验统计量应满足:(1) 可以将样本点估计与零假设值进行比较;(2) 在零假设 H_0 为真时,其分布已知.样本统计量 r 满足这两条原则,它可与零假设值 $\rho=0$ 进行比较,而且在假定 $\rho=0$ 为真时,有已知的抽样分布,由于此分布的概率函数太复杂,在此不赘述,然而 r 的分布关于均值零对称,且 $r=+1$ 或 $r=-1$ 为其尾部的两个极端值.r 的分布是一族分布,对于每个自由度 $v=n-2$,均有唯一的分布,这里 n 为样本中点对的个数.

在零假设 $H_0:\rho=0$ 为真时,利用 r 的抽样分布,对于各种显著性水平 α 和自由度 v,查表可得拒绝零假设 H_0 的临界值.附录中的表 A.11(**Pearson 乘积矩相关系数 r 的临界值**)只给出右侧临界值.由于 r 的分布关于 $r=0$ 对称,故左侧临界值与右侧临界值互为相反数.

记 r 的临界值为 $r_{\alpha/2,v}$(对于双侧假设检验)或 $r_{\alpha,v}$(对于单侧检验),这里 r 是给定显著性水平 α 和自由度 v 时表 A.11 中的值,拒绝零假设 H_0 的临界值决策规则是:

对于双侧检验($H_1:\rho\neq0$),当 $r^*>r_{\alpha/2,v}$ 或 $r^*<-r_{\alpha/2,v}$ 时,拒绝 H_0;

对于右侧检验($H_1:\rho>0$),当 $r^*>r_{\alpha,v}$ 时,拒绝 H_0;

对于左侧检验($H_1:\rho<0$),当 $r^*<-r_{\alpha,v}$ 时,拒绝 H_0.

由表 A.11 可直接得到获得样本值 r^* 的近似概率,它不同于在零假设为真时的总体假定

值 $\rho=0$，只需对给定的 v 在表中找到 r^* 并注意表中所给的概率范围即可.

例 19.10 考虑例 19.8 的大黄蜂试验，已得到 $r=0.626$，给定显著性水平 $\alpha=0.05$，利用表 A.11 中的 r 分布，对零假设 H_0：胸部温度与体重无线性关系，做双侧检验.

解 (1) $H_0:\rho=0$，$H_1:\rho\neq0$；

(2) $\alpha=0.05$；

(3) 对于自由度 $v=n-2=20-2=18$，查表 A.11，可得 r 分布的临界值 r，对于双侧检验，$v=18$，$\alpha=0.05$，查表 A.11 得 $r_{0.05/2,18}=0.444$，因此，决策原则为：当 $r^*<-0.444$ 或 $r^*>0.444$ 时，拒绝 H_0；

(4) 由表 A.11，对 $r^*=0.626$，查得相应 P 值小于 0.01；

(5) 由于 $r^*>0.444$，故拒绝 H_0，接受 H_1. 这一结论可由下述事实确认：在 H_0 为真时，得到 $r^*=0.626$ 的概率小于 0.05. 因此，我们认为：胸部温度与腹部重量之间有显著的线性关系.

19.16 用 t 分布检验关于 ρ 的假设

在 19.15 节中，我们用 r 分布检验了零假设 $H_0:\rho=0$. 对于该检验，亦可用 t 分布，特别当 r 值表中并未提供所需的显著水平时更有用. 为利用 t 分布，r 值应转换成 t 值：

$$t=\frac{r-(\rho=0)}{\sqrt{\frac{1-r^2}{n-2}}} \tag{19.35}$$

这里 r 是样本相关系数，ρ 是总体相关系数(在零假设下为零)，分子为 s_r(相关系数的标准误)，这个统计量服从自由度为 $n-2$ 的 t 分布，表 A.6 给出了它的临界值.

例 19.11 考虑例 19.8 的大黄蜂试验，已知 $r^*=0.626$，给定 $\alpha=0.05$，利用 t 分布，对零假设 H_0：胸部温度与腹部重量之间无线性关系，做双侧检验.

解 (1) $H_0:\rho=0$，$H_1:\rho\neq0$；

(2) $\alpha=0.05$；

(3) 给定自由度 $v=n-2=20-2=18$，及 $\alpha/2=0.05/2$，查表 A.6，对于双侧检验，决策规则为：当 $t^*<-2.101$ 或 $t^*>2.101$ 时，拒绝 H_0；

(4) 将 $r^*=0.626$，$n=20$ 代入(19.35)式，有

$$t^*=\frac{0.626-0}{\sqrt{\frac{1-(0.626)^2}{20-2}}}=3.4058$$

由表 A.6，对于双侧检验，$v=18$，得到 $t^*=3.4058$ 的近似概率为 $0.001<P<0.01$；

(5) 由于 $t^*>2.101$，故拒绝 H_0，接受 H_1. 当零假设为真时，得到 $t^*=3.4058$ 的概率小于 0.05，这一事实亦可确认我们所做的判断. 因而说，胸部温度与腹部重量的相关系数显著不为零.

19.17 用 Z 分布检验假设 $\rho=c$

r 分布和 t 分布均不能用来检验零假设 $H_0:\rho=c$(这里 $c\neq0$ 为给定值). 对于这个检验，可由表 A.10 或数学公式(见 19.14 节)将样本相关系数 r 和假定的总体相关系数 ρ 转化成相应 z 值，分别记为 z_r 和 z_ρ. 当 r 或 ρ 取值于[0,1]时，相应的 z_r 或 z_ρ 取值于 $[0,+\infty)$；当 r 或 ρ 取值于[-1,0]时，相应的 z_r 或 z_ρ 取值于 $(-\infty,0]$，从而可采用下式计算 Z 统计量

$$z=\frac{z_r-z_\rho}{\sqrt{\frac{1}{n-3}}} \tag{19.36}$$

这里 z_r 为样本相关系数 r 的 Z 变换值，z_ρ 为零假设给出的总体相关系数 ρ 的 Z 变换值，分母为估计的标准误，Z 分布的临界值可由表 A.5 得到.

例 19.12 考虑例 19.8 的大黄蜂试验,从另外一个地点取 15 个大黄蜂作为一个样本,我们想知道两次试验的样本是否来自于同一个总体.第二次试验的样本相关系数 $r^*=0.405$,给定 $\alpha=0.05$,对零假设 H_0:总体相关系数 $\rho=0.626$,做双侧检验.

解 ☞ (1) $H_0:\rho=0.626$, $H_1:\rho\neq0.626$;

(2) $\alpha=0.05$;

(3) 当检验相关系数 ρ 是否为 0.626 时,可采用 Z 分布(表 A.5)得到拒绝 H_0 的临界值.因而,决策规则为:当 $z^*<-1.96$ 或 $z^*>1.96$ 时,拒绝 H_0;

(4) 对于 $r^*=0.405$, $\rho=0.626$,由表 A.10,转化成相应 z 值 $z_r=0.4296$, $z_\rho=0.7348$.由于 $n=15$,利用(19.36)式,有

$$z^*=\frac{0.4296-0.7348}{\sqrt{1/12}}=-1.0572\approx-1.06$$

由表 A.5 知,对于标准正态分布,位于区间 $0<z<1.06$ 上方区域的面积为 0.3554.因而,在 H_0 为真时,得到该 z^* 值的概率为

$$2[P(Z>1.06)]=2(0.5000-0.3554)=2(0.1446)=0.2892$$

(5) 由于 $-1.96<z^*<1.96$,故接受 H_0.这一结论可由下述事实得以确认:在 $H_0:\rho=0.626$ 为真时,得到 $r^*=0.405$ 的概率大于 0.05.因此,对于相关系数 ρ 而言,两样本可能来自同一个总体.

19.18 样本相关系数 r 的解释

相本相关系数 r 可通过其平方来解释.统计量 r^2 是由 X, Y 之间的线性关系引起的变差在 X, Y 的总变差中所占比率的一个估计,这种比率一般转化成一个百分数 $(100\%)r^2$,称之为**决定系数**.比如,相关系数 $r=0.4$,表明在样本的总变差中,有 16% 是由 X, Y 之间的线性关系导致的.对于大样本,尽管总变差中由于变量之间的线性关系而产生的那部分变差所占百分比很小,但这个很小的 r 值也可能是显著的.

在所有的统计度量中,相关系数是最常被误解的.一个主要问题是由于因果观念的影响:一个显著相关系数,不论它多么大,都不能被解释成它意味着变量之间的关系是因果关系,这两个变量可能都依赖于第三个变量,而这第三个变量却在我们分析讨论的范围之外.

另外一个问题是当相关系数 r 不是显著不为零时,不能解释成两个变量之间没有关系.这只表明两个变量间无线性关系,它们可以具有非线性关系.

还有另外一种误解:认为相关系数 r 的值是刻画两个变量间线性关系的一种线性尺度.事实上,这是不正确的.比如,$r=0.400$ 与 $r=0.500$ 之间的差异和 $r=0.500$ 与 $r=0.600$ 之间的差异不是同一数量级的.类似地,当比较两个 r 值时,比如,$r=0.8$ 和 $r=0.4$,我们不能说 $r=0.8$ 意味着这种线性关系是 $r=0.4$ 所刻画的线性关系的两倍,我们只能说,$r=0.4$ 表明样本中的总变差的 16% 是由 X, Y 之间的线性关系导致的,而 $r=0.8$ 表明样本中的总变差的 64% 是由 X, Y 之间的线性关系导致的.

此外,当数据是比率或平均值时,可能会产生一些潜在的误差.比率或平均值将变差隐藏在一个观测样本中,这有可能增大相关系数.

例 19.13 用胸部温度和腹部重量之间的线性关系而引起的变差,来描述例 19.8 中的大黄蜂研究结果.

解 ☞ 决定系数 r^2 估计了由于两个变量之间的线性关系而引起的变差的大小,这里 $r^*=0.626$, $r^{*2}=0.3919$,因而我们认为胸部温度与腹部重量的总变差中的 39.2% 是由于它们之间的线性关系而引起的.

19.19 多重相关与偏相关

多重相关度量了三个或三个以上变量之间关系的强度.我们的目的是估计样本的总变差中,由所分析的变量而引起的那部分变差所占比率,由于其计算过程复杂,因而必须借助于计算机.

多重相关系数反映了采用多重相关所分析的所有变量间的总的相关性.若想研究这些变量中某两个变量间的特定关系,不能仅仅同时分析这两个变量,因为这种相关并未包含其它变量对所分析的成对变量的影响.**偏相关**则可用来对多个变量进行成对分析,这种方法考虑了当其他变量视为常量时每对变量之间的相关性.计算机编程再一次成为研究这些关系的最佳工具.

习 题 解 答

最小二乘回归直线

19.1　一位植物学家想根据土壤中的含磷量来预测某种黑麦的长势,她选取了四种含磷量水平:2,4,8,16(百万分之,记为 ppm).在每种磷含量水平条件下种植四株黑麦,当黑麦开花时,测量它们的干重(单位:克),观测数据见表 19.5,求这两个变量之间的最小二乘回归直线.

表 19.5

黑麦	含磷量(ppm)	干重(克)
1	2	4.1
2	2	3.8
3	2	4.0
4	2	3.9
5	4	5.2
6	4	4.9
7	4	5.0
8	4	4.8
9	8	5.7
10	8	5.9
11	8	6.0
12	8	6.2
13	16	11.7
14	16	8.9
15	16	10.1
16	16	10.3

解 ☞　首先,由(19.11)式可得斜率的点估计,表 19.6 给出了所必需的一些和,注意到 $n=16$,将这些和代入(19.11)式,得

表 19.6

	x_i	y_i	x_i^2	y_i^2	x_iy_i
	2	4.1	4	16.81	8.2
	2	3.8	4	14.44	7.6
	2	4.0	4	16.00	8.0
	2	3.9	4	15.21	7.8
	4	5.2	16	27.04	20.8
	4	4.9	16	24.01	19.6
	4	5.0	16	25.00	20.0
	4	4.8	16	23.04	19.2
	8	5.7	64	32.49	45.6
	8	5.9	64	34.81	47.2
	8	6.0	64	36.00	48.0
	8	6.2	64	38.44	49.6
	16	11.7	256	136.89	187.2
	16	8.9	256	79.21	142.4
	16	10.1	256	102.01	161.6
	16	10.3	256	106.09	164.8
$\sum$	120	100.5	1,360	727.49	957.6

$$\hat{b}^* = \frac{957.6 - \frac{(120)(100.5)}{16}}{1.360 - \frac{(120)^2}{16}} = 0.443152 \approx 0.44 \text{ 克/ppm}$$

其次,由(19.6)式可得截距的一个点估计

$$\hat{a}^* = \frac{100.5}{16} - (0.443152)\frac{120}{16} = 2.96 \text{ 克}$$

故最小二乘回归直线为

$$\hat{\mu}^*_{y,x} = 2.96 + 0.44x$$

19.2 一位种群生态学家想确定帽贝的身体大小和总体密度的关系,他在五种不同密度:每平方米300,500,700,900,1100个个体的条件下养殖帽贝,生长五年后,从每个密度组中随机选取四个帽贝,并测量其最大长度(单位:毫米),观测数据见表19.7.求这两个变量之间的最小二乘回归直线?并预测在密度为每平方米600个的条件下,帽贝的生长状况?

表 19.7

帽贝	密度(每平方米的个数)	长度(毫米)
1	300	66
2	300	65
3	300	63
4	300	64
5	500	63
6	500	61
7	500	60
8	500	58
9	700	55
10	700	56
11	700	57
12	700	58
13	900	51
14	900	54
15	900	53
16	900	53
17	1,100	46
18	1,100	48
19	1,100	50
20	1,100	47

解 为求最小二乘回归直线,需先计算 y 的截距 $\hat{a}$ 和斜率 $\hat{b}$ 的点估计,一些所需的和见表19.8.将这些和及 $n=20$ 代入(19.11)式,有

$$\hat{b}^* = \frac{756,000 - \frac{(14,000)(1,128)}{20}}{11,400,000 - \frac{(14,000)^2}{20}} = -0.021 \text{ 毫米/个·米}^2$$

将 $\hat{b}^*$ 代入(19.6)式,得

$$\hat{a}^* = \frac{1,128}{20} - (-0.021)\frac{14,000}{20} = 71.1 \text{ 毫米}$$

于是,最小二乘回归直线为

$$\hat{\mu}^*_{y,x} = 71.1 - 0.021x$$

在每平方米600个的密度条件下预测生长5年后,帽贝的大小为

$$\hat{\mu}_{Y|x=600}=71.1-0.021(600)=58.5\text{ 毫米}$$

表 19.8

	x_i	y_i	x_i^2	y_i^2	x_iy_i
	300	66	90,000	4,356	19,800
	300	65	90,000	4,225	19,500
	300	63	90,000	3,969	18,900
	300	64	90,000	4,096	19,200
	500	63	250,000	3,969	31,500
	500	61	250,000	3,721	30,500
	500	60	250,000	3,600	30,000
	500	58	250,000	3,364	29,000
	700	55	490,000	3,025	38,500
	700	56	490,000	3,136	39,200
	700	57	490,000	3,249	39,900
	700	58	490,000	3,364	40,600
	900	51	810,000	2,601	45,900
	900	54	810,000	2,916	48,600
	900	53	810,000	2,809	47,700
	900	53	810,000	2,809	47,700
	1,100	46	1,210,000	2,116	50,600
	1,100	48	1,210,000	2,304	52,800
	1,100	50	1,210,000	2,500	55,000
	1,100	47	1,210,000	2,209	51,700
$\sum$	14,000	1,128	11,400,000	64,338	756,600

19.3 某比萨饼运输部门的一位经理想说明他们在送货业务上非常有效率,从她的记录中,随机选取了12份定单,送货距离分别为2,5,8,15英里,每个距离送三份订单,对于每个送货服务,记录下比萨饼从商店到顾客所需的时间(单位:分钟),结果数据见表19.9.求时间和距离之间的最小二乘回归直线?如果运送一份比萨饼到恰距商店8英里外的地点,估计需多长时间?

表 19.9

定单	距离(英里)	时间(分钟)
1	2	10.2
2	2	14.6
3	2	18.2
4	5	20.1
5	5	22.4
6	5	30.6
7	10	30.8
8	10	35.4
9	10	50.6
10	15	60.1
11	15	68.4
12	15	72.1

解 首先,由(19.11)式计算斜率的点估计,表19.10给出了所需的一些结果.将这些结果及$n=12$,代入(19.11)式,得

$$\hat{b}=\frac{4,628.5-\dfrac{(96.0)(433.5)}{12}}{1,062.0-\dfrac{(96.0)^2}{12}}=3.9473\approx 3.95\text{ 分钟 / 英里}$$

表 19.10

	x_i	y_i	x_i^2	y_i^2	x_iy_i
	2	10.2	4	104.04	20.4
	2	14.6	4	213.16	29.2
	2	18.2	4	331.24	36.4
	5	20.1	25	404.01	100.5
	5	22.4	25	501.76	112.0
	5	30.6	25	936.36	153.0
	10	30.8	100	948.64	308.0
	10	35.4	100	1,253.16	354.0
	10	50.6	100	2,560.36	506.0
	15	60.1	225	3,612.01	901.5
	15	68.4	225	4,678.56	1,026.0
	15	72.1	225	5,198.41	1,081.5
$\sum$	96	433.5	1,062	20,741.71	4,628.5

由 $\hat{b}^*=3.95$ 及(19.6)式,得截距 $\hat{a}$ 的一个点估计

$$\hat{a}^*=\frac{433.5}{12}-(3.9473)\frac{96.0}{12}=4.55\text{ 分钟}$$

于是最小二乘回归直线为

$$\hat{\mu}^*_{Y|x}=4.55+3.95x$$

由该方程,预测从商店运送一份比萨饼到 8 英里外的地点约需:

$$\mu_{Y|x=8}=4.55+(3.95)(8.0)=36.15\text{ 分钟}$$

方差 $\sigma^2_{Y\cdot X}$ 的估计

19.4 考虑习题 19.1 的植物生长试验,求方差的最佳估计和该估计的标准误?

解 将习题 19.1 中的 $\hat{a}$, $\hat{b}$, n 的值及表 19.6 中的一些结果代入(19.14)式,得方差的估计为

$$s^2_{y\cdot x}=\frac{727.49-(2.96)(100.5)-(0.44)(957.6)}{16-2}=0.619\text{ 克}^2$$

由(19.15)式,得该估计的标准误为

$$s_{y\cdot x}=\sqrt{0.619}=0.7868\approx0.79\text{ 克}$$

19.5 考虑习题 19.2 的帽贝生长试验,求方差的最佳估计和该估计的标准误?

解 将习题 19.2 的 $\hat{a}$, $\hat{b}$, n 的值及表 19.8 中的一些结果代入(19.14)式,得方差的估计为

$$s^2_{y\cdot x}=\frac{64,338-(71.1)(1,128)-(-0.021)(756,600)}{20-2}=1.433\text{ 毫米}^2$$

由(19.15)式,得该估计的标准误为

$$s_{y\cdot x}=\sqrt{1.433}=1.2\text{ 毫米}$$

19.6 考虑习题 19.3 的比萨饼运送试验,求方差的最佳估计和该估计的标准误?

解 将习题 19.3 的 $\hat{a}$, $\hat{b}$, n 的值及表 19.10 中的一些结果代入(19.14)式,得方差的估计为

$$s^2_{y\cdot x}=\frac{20,741.7-(4.55)(433.5)-(3.95)(4,628.5)}{12-2}=48.671\text{ 分钟}^2$$

由(19.15)式,得该估计的标准误为

$$s_{y\cdot x}=\sqrt{48.671}=6.98\text{ 分钟}$$

$\hat{a}$ 和 $\hat{b}$ 的方差

19.7 考虑习题 19.1 的植物生长试验,求 $\hat{a}$ 和 $\hat{b}$ 的点估计的标准差?

解 ☞　将习题 19.4 中的样本值 $s_{y\cdot x}^2=0.619$ 克^2，$n=16$ 及表 19.6 中的一些值代入(19.17)和(19.20)式得 $\hat{a}^*$ 的标准差

$$s_{\hat{a}^*}=\sqrt{s_{\hat{a}^*}^2}=\sqrt{0.619\left[\frac{1}{16}+\frac{\left(\frac{120}{16}\right)^2}{1,360-\frac{(120)^2}{16}}\right]}=0.338\text{ 克}$$

将 $s_{y\cdot x}^2=0.619$ 克^2，$n=16$ 及表 19.6 中的一些值代入(19.19)和(19.21)式得 $\hat{b}^*$ 的标准差

$$s_{\hat{b}^*}=\sqrt{s_{\hat{b}^*}^2}=\sqrt{\frac{0.619}{1,360-\frac{(120)^2}{16}}}=0.037\text{ 克 /ppm}$$

19.8　考虑习题 19.2 的帽贝生长试验，求 $\hat{a}$ 和 $\hat{b}$ 的点估计的标准差？

解 ☞　将习题 19.5 中的样本值 $s_{y\cdot x}^2=1.433$ 毫米^2，$n=20$ 及表 19.8 中的一些值代入(19.17)和(19.20)式得 $\hat{a}^*$ 的标准差

$$s_{\hat{a}^*}=\sqrt{s_{\hat{a}^*}^2}=\sqrt{1.433\left[\frac{1}{20}+\frac{\left(\frac{14,000}{20}\right)^2}{11,400,000-\frac{(14,000)^2}{20}}\right]}=0.714\text{ 毫米}$$

将 $s_{y\cdot x}^2=1.433$ 毫米^2、$n=20$ 及表 19.8 中的一些值代入(19.19)和(19.21)式得 $\hat{b}^*$ 的标准差

$$s_{\hat{b}^*}=\sqrt{s_{\hat{b}^*}^2}=\sqrt{\frac{1.433}{11,400,000-\frac{(14,000)^2}{20}}}=0.00095\text{ 毫米 / 个·米}^2$$

19.9　考虑习题 19.3 的比萨饼运送试验，求 $\hat{a}$ 和 $\hat{b}$ 的点估计的标准差？

解 ☞　将习题 19.6 中的样本值 $s_{y\cdot x}^2=48.671$ 分钟^2，$n=12$ 及表 19.10 中的一些值代入(19.17)和(19.20)式得 $\hat{a}^*$ 的标准差

$$s_{\hat{a}^*}=\sqrt{s_{\hat{a}^*}^2}=\sqrt{48.671\left[\frac{1}{12}+\frac{\left(\frac{96}{12}\right)^2}{1,062-\frac{(96)^2}{12}}\right]}=3.828\text{ 分钟}$$

将 $s_{y\cdot x}^2=48.671$ 分钟^2，$n=12$ 及表 19.10 中的一些值代入(19.19)和(19.21)式得 $\hat{b}^*$ 的标准差

$$s_{\hat{b}^*}=\sqrt{s_{\hat{b}^*}^2}=\sqrt{\frac{48.671}{1,062-\frac{(96)^2}{12}}}=0.407\text{ 分钟 / 英里}$$

a 和 *b* 的置信区间

19.10　考虑习题 19.1 的植物生长试验，求 a 和 b 的 95% 置信区间？

解 ☞　由习题 19.1 知 $\hat{a}^*=2.96$ 克，$n=16$，由习题 19.7 知 $s_{\hat{a}^*}=0.338$ 克，由表 A.6 得临界值 $t_{\alpha/2,\nu}=t_{0.05/2,14}=2.145$。于是，由(19.23)式可得 a 的置信区间为

$$2.96\text{ 克}\pm(2.145\times0.338\text{ 克})$$

$$2.96\text{ 克}\pm0.725\text{ 克}$$

由习题 19.1 知 $\hat{b}^*=0.44$ 克/ppm，$n=16$，由习题 19.7 知 $s_{\hat{b}^*}=0.037$ 克/ppm，由表 A.6 得临界值 $t_{\alpha/2,\nu}=t_{0.05/2,14}=2.145$。于是，由(19.25)式可得 b 的置信区间为

$$0.44\text{ 克 /ppm}\pm(2.145\times0.037\text{ 克 /ppm})$$

$$0.44\text{ 克 /ppm}\pm0.079\text{ 克 /ppm}$$

19.11　考虑习题 19.2 的帽贝生长试验，求 a 和 b 的 99% 置信区间？

解 ☞　由习题 19.2 知 $\hat{a}^*=71.1$ 毫米，$n=20$，由习题 19.8 知 $s_{\hat{a}^*}=0.714$ 毫米，由表 A.6 得临界值 $t_{\alpha/2,\nu}=t_{0.01/2,18}=2.878$。于是，由(19.23)式可得 a 的置信区间为

$$71.1\text{毫米} \pm (2.878 \times 0.714\text{毫米})$$

$$71.1\text{毫米} \pm 2.05\text{毫米}$$

由习题 19.2 知 $\hat{b}^* = -0.021$，$n=20$，由习题 19.8 知 $s_{b^*} = 0.00095$，由表 A.5 得临界值 $t_{\alpha/2,\nu} = t_{0.01/2,18} = 2.878$. 于是，由(19.25)式可得 b 的置信区间为

$$-0.021\text{毫米/个·米}^2 \pm (2.878 \times 0.00095\text{毫米/个·米}^2)$$

$$-0.021\text{毫米/个·米}^2 \pm 0.0027\text{毫米/个·米}^2$$

19.12　考虑习题 19.3 的比萨饼运送试验，求 a 和 b 的 90%置信区间？

解 ☞　由习题 19.3 知 $\hat{a}^* = 4.55$ 分钟，$n=12$，由习题 19.9 知 $s_{a^*} = 3.828$ 分钟，由表 A.6 得临界值 $t_{\alpha/2,\nu} = t_{0.10/2,10} = 1.812$. 于是，由(19.23)式可得 a 的置信区间为

$$4.55\text{分钟} \pm (1.812 \times 3.828\text{分钟})$$

$$4.55\text{分钟} \pm 6.936\text{分钟}$$

由习题 19.3 知 $\hat{b}^* = 3.95$ 分钟/英里，由习题 19.9 知 $s_{b^*} = 0.407$ 分钟/英里，于是，由(19.25)式可得 b 的置信区间为

$$3.95\text{分钟/英里} \pm (1.812 \times 0.407\text{分钟/英里})$$

$$3.95\text{分钟/英里} \pm 0.737\text{分钟/英里}$$

方差 $\sigma^2_{Y\cdot X}$ 的置信区间

19.13　考虑习题 19.1 的植物生长试验，求植物生长与含磷量之间的线性关系的总体方差 $\sigma^2_{Y\cdot X}$ 的 95%置信区间？

解 ☞　由习题 19.1 知 $n=16$，由习题 19.4 知 $s^2_{Y\cdot X} = 0.619$ 克^2. 由表 A.7，给定 $\nu = n-2 = 14$，$\chi^2_{0.025,14} = 26.12$，$\chi^2_{0.975,14} = 5.63$，利用(19.27)式，可得置信区间

$$\left[\frac{(14 \times 0.619\text{克}^2)}{26.12}, \frac{(14 \times 0.619\text{克}^2)}{5.63}\right]$$

$$(0.3318\text{克}^2, 1.5393\text{克}^2)$$

19.14　考虑习题 19.2 的帽贝生长试验，求帽贝的大小与总体密度之间的线性关系的总体方差 $\sigma^2_{Y\cdot X}$ 的 99%置信区间？

解 ☞　由习题 19.2 知 $n=20$，由习题 19.5 知 $s^2_{Y\cdot X} = 1.433$ 毫米^2. 由表 A.7，给定 $\nu = n-2 = 18$，$\chi^2_{0.005,18} = 37.16$，$\chi^2_{0.995,18} = 6.26$，利用(19.27)式，可得置信区间

$$\left[\frac{(18 \times 1.433\text{毫米}^2)}{37.16}, \frac{(18 \times 1.433\text{毫米}^2)}{6.26}\right]$$

$$(0.6941\text{毫米}^2, 4.1204\text{毫米}^2)$$

19.15　考虑习题 19.3 的比萨饼运送试验，求运送时间与运送距离之间线性关系的总体方差 $\sigma^2_{Y\cdot X}$ 的 90%置信区间？

解 ☞　由习题 19.3 知 $n=12$，由习题 19.6 知 $s^2_{Y\cdot X} = 48.671$ 分钟^2. 由表 A.7，给定 $\nu = n-2 = 10$，$\chi^2_{0.05,10} = 18.31$，$\chi^2_{0.95,10} = 3.94$，利用(19.27)式，可得置信区间

$$\left[\frac{(10 \times 48.671\text{分钟}^2)}{18.31}, \frac{(10 \times 48.671\text{分钟}^2)}{3.94}\right]$$

$$(26.582\text{分钟}^2, 123.530\text{分钟}^2)$$

Y 的期望的预测区间

19.16　在含磷量为 5ppm 的土壤上种植黑麦，试利用习题 19.1 确定的最小二乘回归方程预测一株黑麦的重量，并求其 95%的预测区间？

解 ☞　由习题 19.1 知回归方程为 $\hat{\mu}^*_{Y\cdot x} = 2.96 + 0.44x$. 则种植在含磷量 5ppm 的土壤中的一

株黑麦的预测重量为：$\hat{\mu}_{Y|x} = 2.96 + (0.44 \times 5) = 5.16$ 克．由于 $n = 16 < 30$，故采用 t 分布求预测区间．首先利用(19.29)式求 Y 的均值的标准误．由习题 19.4 知 $s_{y \cdot x} = 0.79$ 克．利用 $x_0 = 5$ppm，(19.8)式及表 19.6 给出的一些值，我们有

$$s_{\hat{\mu}_{Y|x=5}} = 0.79\sqrt{1 + \frac{1}{16} + \frac{\left(5 - \frac{120}{16}\right)^2}{1,360 - \frac{(120)^2}{16}}} = 0.820 \text{ 克}$$

由表 A.6 得临界值 $t_{0.05/2,14} = 2.145$，且 $\hat{\mu}_{Y|x=5} = 5.16$ 克，$s_{\hat{\mu}_{Y|x}} = 0.820$ 克，故预测区间为

$$5.16 \text{ 克} \pm (2.145 \times 0.820 \text{ 克})$$

$$5.16 \text{ 克} \pm 1.759 \text{ 克}$$

19.17　已知某海边帽贝的密度为 850 个/米^2，试利用习题 19.2 确定的最小二乘回归直线预测这些帽贝的平均长度，并求其 99% 预测区间？

解 ☞　由习题 19.2 知回归方程为 $\hat{\mu}_{Y|x} = 71.1 - 0.021x$，所以生活在总体密度为 850 个/米^2 的海边的帽贝的预测长度为 $\hat{\mu}_{Y|x=850} = 71.1 - (0.021 \times 850) = 53.2$ 毫米．由于 $n = 20 < 30$，故利用 t 分布求预测区间．首先利用(19.29)式求出 Y 值的均值的标准差．由习题 19.5 知 $s_{y \cdot x} = 1.2$ 毫米；而 $x_0 = 850$，利用(19.8)式及表 19.8 中的一些值，得

$$s_{\hat{\mu}_{Y|x=850}} = 1.2\sqrt{1 + \frac{1}{20} + \frac{\left(850 - \frac{14,000}{20}\right)^2}{11,400,000 - \frac{(14,000)^2}{20}}} = 1.24 \text{ 毫米}$$

查表 A.6 得临界值 $t_{0.01/2,18} = 2.878$，又 $\hat{\mu}_{Y|x=850} = 53.2$ 毫米，$s_{\hat{\mu}_{Y|x=850}} = 1.24$ 毫米，故预测区间为

$$53.2 \text{ 毫米} \pm (2.878 \times 1.24 \text{ 毫米})$$

$$53.2 \text{ 毫米} \pm 3.57 \text{ 毫米}$$

19.18　试利用习题 19.3 确定的最小二乘回归直线，预测该比萨饼店运送一份比萨饼到 4.8 英里处地点所需时间，并求其 90% 预测区间？

解 ☞　由习题 19.3 知回归方程为 $\hat{\mu}_{Y|x} = 4.55 + 3.95x$．一份比萨饼由商店运送到 4.8 英里处某地的预测时间为：$\hat{\mu}_{Y|x=4.8} = 4.55 + (3.95 \times 4.8) = 23.51$ 分钟．由于 $n = 12 < 30$，故利用 t 分布求预测区间．首先，利用(19.29)式求 Y 的均值的标准误．由习题 19.6 得 $s_{y \cdot x} = 6.98$ 分钟．这里 $x_0 = 4.8$ 英里，利用(19.8)式及表 19.10 中的一些值，得

$$s_{\hat{\mu}_{Y|x=4.8}} = 1.2\sqrt{1 + \frac{1}{12} + \frac{\left(4.8 - \frac{96}{12}\right)^2}{1,062.0 - \frac{(96)^2}{12}}} = 7.381 \text{ 分钟}$$

查表 A.6 得临界值 $t_{0.10/2,10} = 1.812$，又 $\hat{\mu}_{Y|x=4.8} = 23.51$ 分钟，$s_{\hat{\mu}_{Y|x=4.8}} = 7.381$ 分钟，故预测区间为

$$23.51 \text{ 分钟} \pm (1.812 \times 7.381 \text{ 分钟})$$

$$23.51 \text{ 分钟} \pm 13.374 \text{ 分钟}$$

关于斜率 b 的假设检验

19.19　在习题 19.1 中，我们求得最小二乘回归直线的斜率为 $\hat{b}^* = 0.44$ 克/ppm，由习题 19.7 得 $\hat{b}$ 的标准差为 0.037 克/ppm．给定 $\alpha = 0.01$，利用临界值决策规则，对零假设 $H_0: b = 0$，做右侧检验．

解 ☞

(1) $H_0: b = 0$，$H_1: b > 0$；

(2) $\alpha = 0.01$；

(3) 我们利用(19.24)式给出的检验统计量．由表 A.6 知，对于给定 $\alpha = 0.01$，$v = n - 2 = 14$ 的

右侧检验,临界值为 2.624,决策规则为:若 $t^* > 2.624$,则拒绝 H_0.

(4) 将 $\hat{b}^* = 0.44$ 克/ppm,$s_{b^*} = 0.037$ 克/ppm,代入(19.24)式,得

$$t^* = \frac{0.44 - 0}{0.037} = 11.892$$

查表 A.6,对于 $v = 14$,得到该样本值的概率为 $P < 0.0005$.

(5) 由于 $t^* > 2.624$,故拒绝 H_0,接受 H_1.这一结论可由下述事实确认:在零假设 H_0 为真时,得到 $t^* = 11.892$ 的概率小于 0.01.因而,黑麦重量与含磷量之间关系的斜率大于零.

19.20 在习题 19.2 中,我们求得最小二乘回归直线的斜率为 $\hat{b}^* = -0.021$ 毫米/个·米^2,又由习题 19.8 知 $\hat{b}$ 的标准差为 0.00095 毫米/个·米^2.给定 $\alpha = 0.05$,利用临界值决策规则,对零假设 $H_0: b = 0$,做左侧检验.

解 ☞

(1) $H_0: b = 0$, $H_1: b < 0$;

(2) $\alpha = 0.05$;

(3) 我们利用(19.24)式定义的检验统计量.由表 A.6 知,对于 $\alpha = 0.05$,$v = n - 2 = 18$ 的左侧检验,临界值为 1.734,决策规则为:若 $t^* < -1.734$,则拒绝 H_0;

(4) 将 $\hat{b}^* = -0.021$ 毫米/个·米^2,$s_{b^*} = 0.00095$ 毫米/个·米^2 代入(19.24)式有

$$t^* = \frac{-0.021 - 0}{0.00095} = -22.105$$

查表 A.6,对于 $v = 18$,得到该 t^* 的近似概率为 $P < 0.0005$.

(5) 由于 $t^* < -1.734$,故拒绝 H_0,接受 H_1.这一结论可由下述事实确认:在 H_0 为真时,得到 $t^* = -22.105$ 的概率小于 0.05.因而帽贝长度与总体密度之间关系的斜率小于零.

19.21 在习题 19.3 中,我们得到最小二乘回归直线的斜率为 $\hat{b}^* = 3.95$ 分钟/英里,又由习题 19.9 知 $\hat{b}$ 的标准差为 0.407 分钟/英里.给定 $\alpha = 0.05$,利用临界值决策规则,对零假设 $H_0: b = 0$,做右侧检验

解 ☞

(1) $H_0: b = 0$, $H_1: b > 0$;

(2) $\alpha = 0.05$;

(3) 我们利用(19.24)式定义的检验统计量.由表 A.6 知,对于 $\alpha = 0.05$,$v = n - 2 = 10$ 的右侧检验,临界值为 1.812,决策规则为:若 $t^* > 1.812$,则拒绝 H_0;

(4) 将 $\hat{b}^* = 3.95$ 分钟/英里和 $s_{b^*} = 0.407$ 分钟/英里代入(19.24)式有

$$t^* = \frac{3.95 - 0}{0.407} = 9.705$$

查表 A.6,对于 $v = 10$,得到该 t^* 的近似概率为 $P < 0.0005$.

(5) 由于 $t^* > 1.812$,故拒绝 H_0,接受 H_1.这一结论可由下述事实确认:在 H_0 为真时,得到 $t^* = 9.705$ 的概率小于 0.05.因而运送时间与运送距离之间关系的斜率大于零.

样本相关系数 r 的计算

19.22 鸟类通过急促喘息散发余热,我们想研究鸟的体温(℃)与呼吸速率是否有线性关系.在不同的环境温度下,随机抽取 15 只鸟,对每只鸟测量它的体温和每分钟的呼吸次数,测量结果见表 19.11,计算这两个变量间的相关系数.

解 ☞　由于环境温度与鸟类每分钟的呼吸次数可以假定服从二元正态分布,因而可以用样本相关系数 r 来估计总体相关系数 ρ.为计算样本相关系数 r,我们需先计算 $\sum x$, $\sum x^2$, $\sum y$, $\sum y^2$ 和 $\sum xy$,这些值见表 19.12.将这些值代入(19.32)式有

$$r^* = \frac{32,733.5 - \left[\frac{(603.5)(804)}{15}\right]}{\sqrt{(24,301.79) - \frac{(603.5)^2}{15}}\sqrt{(52,446) - \frac{(804)^2}{15}}} = 0.871$$

样本相关系数 $r^* = 0.871$ 表明环境温度与鸟类的呼吸速率有正向的线性关系.

表 19.11

鸟	体温(℃)	呼吸速率(每分钟呼吸次数)
1	39.6	33
2	40.1	50
3	41.7	75
4	39.0	13
5	41.9	68
6	42.8	115
7	40.3	52
8	39.0	33
9	39.7	60
10	39.3	28
11	41.8	74
12	39.6	39
13	40.2	76
14	39.1	30
15	39.4	58

表 19.12

	x_i	y_i	x_i^2	y_i^2	x_iy_i
	39.6	33	1,568.16	1,089	1,306.8
	40.1	50	1,608.01	2,500	2,005.0
	41.7	75	1,738.89	5,625	3,127.5
	39.0	13	1,521.00	169	507.0
	41.9	68	1,755.61	4,624	2,849.2
	42.8	115	1,831.84	13,225	4,922.0
	40.3	52	1,624.09	2,704	2,095.6
	39.0	33	1,521.00	1,089	1,287.0
	39.7	60	1,576.09	3,600	2,382.0
	39.3	28	1,544.49	784	1,100.4
	41.8	74	1,747.24	5,476	3,093.2
	39.6	39	1,568.16	1,521	1,544.4
	40.2	76	1,616.04	5,776	3,055.2
	39.1	30	1,528.81	900	1,173.0
	39.4	58	1,552.36	3,364	2,285.2
$\sum$	603.5	804	24,301.79	52,446	32,733.5

19.23 一位生物地理学家认为一个岛屿自脱离大陆以来,生活在该岛上的一些物种正在逐渐减少.假定该岛屿完全脱离大陆的时刻为该岛屿年龄起始时刻,为研究岛屿的年龄与生活在该岛上的蜥蜴种类数是否有线性关系,她随机抽取 13 个岛屿,计算出每个岛屿的年龄及生活在该岛上的蜥蜴种类数,结果如表 19.13 所示.计算岛屿年龄与蜥蜴种类数的相关系数.

表 19.13

岛屿年龄(年)	蜥蜴种类(种类数量)
9,100	75
4,500	55
13,400	12
6,500	53
11,200	48
12,000	26
5,100	50
7,400	55
10,600	21
11,800	48
14,700	32
8,800	25
10,900	21

解 ☞　由于岛屿年龄与物种数量可以假定服从二元正态分布,因而可以用样本相关系数 r 来估计总体相关系数ρ. 为计算样本相关系数 r,我们需先计算 $\sum x$, $\sum x^2$, $\sum y$, $\sum y^2$ 和 $\sum xy$,这些值见表 19.14.

表 19.14

	x_i	y_i	x_i^2	y_i^2	x_iy_i
	9,100	75	82,810,000	5,625	682,500
	4,500	55	20,250,000	3,025	247,500
	13,400	12	179,560,000	144	160,800
	6,500	53	42,250,000	2,809	344,500
	11,200	48	125,440,000	2,304	537,600
	12,000	26	144,000,000	676	312,000
	5,100	50	26,010,000	2,500	255,000
	7,400	55	54,760,000	3,025	407,000
	10,600	21	112,360,000	441	222,600
	11,800	48	139,240,000	2,304	566,400
	14,700	32	216,090,000	1,024	470,400
	8,800	25	77,440,000	625	220,000
	10,900	21	118,810,000	441	228,900
Σ	126,000	521	1,339,020,000	24,943	4,655,200

将这些值代入(19.32)式有

$$r^* = \frac{4,655,200 - \left[\frac{(126,000)(521)}{13}\right]}{\sqrt{(1,339,020,000) - \frac{(126,000)^2}{13}}\sqrt{(24,943) - \frac{(521)^2}{13}}} = 0.570$$

样本相关系数 $r^* = -0.570$ 表明岛屿年龄与蜥蜴种类的数量有负向的线性关系.

19.24　某公司经理欲向应聘者保证雇员的初始年薪与工作 10 年后的年薪无关. 为确认这一点,他随机抽取 12 个工作多年的雇员,记录下他们的初始年薪和工作 10 年后的年薪,结果见表 19.15. 计算初始年薪和工作 10 年后的年薪的相关系数.

19.15

雇员	初始年薪	10 年后年薪
1	$ 26,000	$ 37,000
2	$ 42,000	$ 90,000
3	$ 37,000	$ 48,000
4	$ 82,000	$ 90,000
5	$ 66,000	$ 88,000
6	$ 44,000	$ 100,000
7	$ 24,000	$ 95,000
8	$ 39,000	$ 120,000
9	$ 55,000	$ 95,000
10	$ 61,000	$ 76,000
11	$ 77,000	$ 89,000
12	$ 58,000	$ 100,000

解 ☞　由于公司内的这两种年薪可以假定服从二元正态分布,因而可以用样本相关系数 r 来估计总体相关系数ρ. 为计算样本相关系数 r,我们需先计算 $\sum x$, $\sum x^2$, $\sum y$, $\sum y^2$ 和 $\sum xy$,这些值见表 19.16 .

将这些值代入(19.32)式有

$$r^* = \frac{53,580,000,000 - \left[\frac{(611,000)(1,028,000)}{12}\right]}{\sqrt{(34,961,000,000) - \frac{(611,000)^2}{12}}\sqrt{(93,764,000,000) - \frac{(1,028,000)^2}{12}}}$$

$= 0.264$

样本相关系数 $r^* = 0.264$ 较小表明两种年薪之间没有线性关系.

表 19.16

	x_i	y_i	x_i^2	y_i^2	x_iy_i
	26,000	37,000	676,000,000	1,369,000,000	962,000,000
	42,000	90,000	1,764,000,000	8,100,000,000	3,780,000,000
	37,000	48,000	1,369,000,000	2,304,000,000	1,776,000,000
	82,000	90,000	6,724,000,000	8,100,000,000	7,380,000,000
	66,000	88,000	4,356,000,000	7,744,000,000	5,808,000,000
	44,000	100,000	1,936,000,000	10,000,000,000	4,400,000,000
	24,000	95,000	576,000,000	9,025,000,000	2,280,000,000
	39,000	120,000	1,521,000,000	14,400,000,000	4,680,000,000
	55,000	95,000	3,025,000,000	9,025,000,000	5,225,000,000
	61,000	76,000	3,721,000,000	5,776,000,000	4,636,000,000
	77,000	89,000	5,929,000,000	7,921,000,000	6,853,000,000
	58,000	100,000	3,364,000,000	10,000,000,000	5,800,000,000
$\sum$	611,000	1,028,000	34,961,000,000	93,764,000,000	53,580,000,000

总体相关系数 ρ 的置信区间

19.25 考虑习题 19.22 中鸟的体温与呼吸速率的关系，求总体相关系数 ρ 的 95% 置信区间?

解 首先，将样本相关系数 r 转化成相应的 z_r 值. 由表 A.10 知，对于 $r^* = 0.871$，$z_r = 1.3372$，因而置信限为(见 19.14 节)

$$[1.3372 - z_{\alpha/2}\sigma_{z_r}, 1.3372 + z_{\alpha/2}\sigma_{z_r}]$$

为求 95% 置信区间. 由表 A.5 知 $z_{0.05/2} = z_{0.025} = 1.96$，又 $n = 15$，代入(19.33)式得标准误

$$\sigma_{z_r} = \sqrt{\frac{1}{n-3}} = \sqrt{\frac{1}{12}} = 0.2887$$

置信限为

$$[1.3372 - (1.96 \times 0.2887),\quad 1.3372 + (1.96 \times 0.2887)]$$

$$(0.7713,\quad 1.9031)$$

再由表 A.10，将这些转化成 r 值，这样得 r^* 的 95% 置信区间为(0.648, 0.956)，因而以 95% 置信度，我们说总体相关系数 ρ 位于区间(0.648, 0.956)内.

19.26 考虑习题 19.23 中岛屿年龄与蜥蜴种类数的关系，求总体相关系数 ρ 的 95% 置信区间.

解 首先，将样本相关系数 r 转化成相应的 z_r 值. 由表 A.10 知，对于 $r^* = -0.570$，$z_r = -0.6475$，因而置信限为(见 19.14 节)

$$[-0.6475 - z_{\alpha/2}\sigma_{z_r}, -0.6475 + z_{\alpha/2}\sigma_{z_r}]$$

为求 95% 置信区间，由表 A.5 知 $z_{0.05/2} = z_{0.025} = 1.96$，又 $n = 13$，代入(19.33)式得标准误

$$\sigma_{z_r} = \sqrt{\frac{1}{n-3}} = \sqrt{\frac{1}{10}} = 0.3162$$

置信限为

$$[-0.6475 - (1.96 \times 0.3162),\quad 0.6475 + (1.96 \times 0.3162)]$$

$$(-1.2673, -0.0277)$$

再由表 A.10，将这些转化成 r 值，从而得总体相关系数 ρ 的 95% 置信区间为(−0.853, −0.028).

19.27 考虑习题 19.24 中雇员初始年薪与工作 10 年后年薪之间的关系，求总体相关系数 ρ 的 95% 置信区间.

解 ☞ 首先,将样本相关系数 r 转化成相应的 z_r 值.由表 A.10 知,对于 $r^*=0.264$, $z_r=0.2704$.因而置信限为(见 19.14 节)

$$[0.2704-z_{\alpha/2}\sigma_{z_r},\ 0.2704+z_{\alpha/2}\sigma_{z_r}]$$

为求 95% 置信区间,由表 A.5 知 $z_{0.05/2}=z_{0.025}=1.96$,又 $n=12$,代入(19.33)式得标准误

$$\sigma_{z_r}=\sqrt{\frac{1}{n-3}}=\sqrt{\frac{1}{9}}=0.3333$$

置信限为

$$[0.2704-(1.96\times 0.3333),\ 0.2704+(1.96\times 0.3333)]$$
$$(-0.3829,\ 0.9237)$$

再由表 A.10,将这些转化成 r 值,从而得总体相关系数 ρ 的 95% 置信区间为$(-0.365,\ 0.728)$.

用 r 分布检验关于相关系数 ρ 的假设

19.28 假定我们在做习题 19.22 中的试验之前,认为鸟类呼吸速率随体温升高而增大.给定 $\alpha=0.01$,利用临界值决策规则,对零假设 $H_0:\rho=0$ 和 $H_1:\rho>0$,做右侧检验(这里假定 $r^*=0.871$).

解 ☞

(1) $H_0:\rho=0$, $H_1:\rho>0$;

(2) $\alpha=0.01$;

(3) 由表 A.11 知,对于 $\alpha=0.01$, $v=13$ 的右侧检验,临界值为 $r_{0.01,13}=0.592$.因而,决策规则为:若 $r^*>0.592$,则拒绝 H_0;

(4) 查表 A.11 知,对于 $v=13$ 的右侧检验,得到 $r^*=0.871$ 的近似概率小于 0.005;

(5) 由于 $r^*>0.592$,故拒绝 H_0,接受 H_1.这一结论可由下述事实确认:在 $H_0:\rho=0$ 为真时,得到样本值 $r^*=0.871$ 的概率小于 0.01,因而鸟的体温与呼吸速率之间有显著的线性关系.

19.29 考虑习题 19.23 中岛屿的年龄和蜥蜴种类数的关系.已知 $r^*=-0.570$,给定 $\alpha=0.05$,利用临界值决策规则,对零假设 $H_0:\rho=0$,做双侧检验.

解 ☞

(1) $H_0:\rho=0$, $H_1:\rho\neq 0$;

(2) $\alpha=0.05$;

(3) 由表 A.11 知,对于 $\alpha=0.05$, $v=11$ 的双侧检验,临界值为 $r_{0.05/2,11}=0.553$.因而,决策规则为:若 $r^*<-0.553$ 或 $r^*>0.553$,则拒绝 H_0;

(4) 查表 A.11 知,对于 $v=11$ 的双侧检验,得到 $r^*=0.570$ 的近似概率为 $0.02<P<0.05$;

(5) 由于 $r^*<-0.553$,故拒绝 H_0,接受 H_1.这一结论可由下述事实确认:在 $H_0:\rho=0$ 为真时,得到样本值 $r^*=-0.570$ 的概率小于 0.05,因而岛屿年龄与蜥蜴种类数之间有显著的反向线性关系.

19.30 考虑习题 19.24 中雇员的初始年薪与工作 10 年后年薪之间的关系.已知 $r^*=0.264$,给定 $\alpha=0.05$,利用临界值决策规则,对零假设 $H_0:\rho=0$,做双侧检验.

解 ☞

(1) $H_0:\rho=0$, $H_1:\rho\neq 0$;

(2) $\alpha=0.05$;

(3) 由表 A.11 知,对于 $\alpha=0.05$, $v=10$ 的双侧检验,临界值为 $r_{0.05/2,10}=0.576$.因而,决策规则为:若 $r^*<-0.576$ 或 $r^*>0.576$,则拒绝 H_0;

(4) 查表 A.11 知,对于 $v=10$ 的双侧检验,得到 $r^*=0.264$ 的近似概率为 $P>0.10$;

(5) 由于 $-0.576<r^*<0.576$,故接受 H_0.这一结论可由下述事实确认:在 $H_0:\rho=0$ 为真时,得到样本值 $r^*=0.264$ 的概率大于 0.05,因而这两种年薪之间线性关系不显著.

用 t 分布检验关于 ρ 的假设

19.31 给定 $\alpha=0.01$,用 t 变换和 t 分布表,利用临界值决策规则,对习题 19.28 中的零假设

H_0:鸟类的呼吸速率与体温之间没有线性关系,做单侧检验.

解 ☞

(1) $H_0:\rho=0$, $H_1:\rho>0$;

(2) $\alpha=0.01$;

(3) 利用表A.6,对于 $\alpha=0.01$, $v=n-2=15-2=13$ 的单侧检验,决策规则为:若 $t^*>2.650$,则拒绝 H_0.

(4) 由习题19.22知 $r^*=0.871$,将之代入(19.35)式得

$$t^*=\frac{0.871-0}{\sqrt{\dfrac{1-(0.871)^2}{15-2}}}=6.392$$

查表A.6知,对于 $v=13$ 的单侧检验,得到该 t^* 值的近似概率为 $P<0.0005$.

(5) 由于 $t^*>2.650$,故拒绝 H_0,接受 H_1.这一结论可由下述事实确认:在 H_0 为真时,得到 $t^*=6.392$ 的近似概率小于0.01.因而,鸟类的体温与呼吸速率之间有显著的正向相关.

19.32 给定 $\alpha=0.05$,用 t 变换和 t 分布表,利用临界值决策规则,对习题19.29中的零假设 H_0:岛屿年龄与蜥蜴种类的数量无线性关系,做双侧检验.

解 ☞

(1) $H_0:\rho=0$, $H_1:\rho\neq0$;

(2) $\alpha=0.05$;

(3) 利用表A.6,对于 $\alpha=0.05$, $v=n-2=13-2=11$ 的双侧检验,决策规则为:若 $t^*>2.201$ 或 $t^*<-2.201$,则拒绝 H_0;

(4) 由习题19.22知 $r^*=-0.570$,将之代入(19.35)式得

$$t^*=\frac{-0.570-0}{\sqrt{\dfrac{1-(-0.570)^2}{13-2}}}=-2.301$$

查表A.6知,对于 $v=11$ 的双侧检验,得到 $t^*=-2.301$ 的近似概率为 $0.02<P<0.05$;

(5) 由于 $t^*<-2.201$,故拒绝 H_0,接受 H_1.这一结论可由下述事实确认:在 $H_0:\rho=0$ 为真时,得到 $t^*=-2.301$ 的近似概率小于0.05.因而,岛屿的年龄与蜥蜴种类的数量之间有反向的线性关系.

19.33 给定 $\alpha=0.05$,用 t 变换和 t 分布表,利用临界值决策规则,对习题19.30中的零假设 H_0:雇员的初始年薪与工作10后的年薪无线性关系,做双侧检验.

解 ☞

(1) $H_0:\rho=0$, $H_1:\rho\neq0$;

(2) $\alpha=0.05$;

(3) 利用表A.6,对于 $\alpha=0.05$, $v=n-2=12-2=10$ 的双侧检验,决策原则为:若 $t^*>2.228$ 或 $t^*<-2.228$,则拒绝 H_0;

(4) 由习题19.22知 $r^*=0.264$,将之代入(19.35)式得

$$t^*=\frac{0.264-0}{\sqrt{\dfrac{1-(0.264)^2}{12-2}}}=0.866$$

查表A.6知,对于 $v=10$ 的双侧检验,得到 $t^*=0.264$ 的近似概率为 $P>0.20$;

(5) 由于 $-2.228<t^*<2.228$,故接受 H_0.这一结论可由下述事实确认:在 H_0 为真时,得到 $t^*=0.866$ 的近似概率大于0.05.因而,两种年薪之间有显著的线性关系.

用 Z 分布检验假设 $\rho=c$

19.34 重新考虑习题19.22中的关于鸟类体温和呼吸速率的关系的研究,这次用来自另外一种鸟类的15只鸟做该试验,并得到样本相关系数 $r^*=0.600$,我们想知道分别由两个样本得到的鸟类体温与呼吸速率的关系是否相同.给定 $\alpha=0.05$,利用临界值决策规

则,对零假设 H_0:总体相关系数 $\rho=0.871$,做双侧检验.

解 ☞

(1) $H_0:\rho=0.871$, $H_1:\rho\neq0.871$;

(2) $\alpha=0.05$;

(3) 由于是为了检验相关系数是否等于某个特定值,故采用 Z 分布(表 A.5)来寻找临界值.查表 A.5 得 $z_{0.05/2}=1.96$,决策规则为:若 $z^*<-1.96$ 或 $z^*>1.96$,则拒绝 H_0;

(4) 利用表 A.10 将 $r=0.600$ 和 $\rho=0.871$ 转化成相应的 z 值,有 $z_r=0.6931$, $z_\rho=1.3372$.将这些 z 值及 $n=15$ 代入(19.36)式,得 z 的样本值

$$z^*=\frac{0.6931-1.3372}{\sqrt{1/12}}=-2.2312$$

由表 A.5 知,对于标准正态分布,位于区间 $0\leqslant z\leqslant2.23$ 上方区域的面积为 0.4871 因而在 H_0 为真时,得到该 z 值的概率为

$$2[P(Z>2.23)]=2(0.5000-0.4871)=2(0.0129)=0.0258$$

(5) 由于 $z^*<-1.96$,故拒绝 H_0,接受 H_1.这一结论可由下述事实确认:在 $H_0:\rho=0.871$ 为真时,得到 $r^*=0.600$ 的概率小于 0.05.因而从总体相关系数 ρ 考虑,不像是两个样本来自同一个总体.

19.35 重新考虑习题 19.23 中的关于岛屿年龄与蜥蜴种类数的关系的研究,这次随机抽取 20 个岛屿,计算得样本相关系数 $r^*=-0.410$.给定 $\alpha=0.05$,利用临界值决策规则,对假设 H_0:总体相关系数 $\rho=-0.570$ 和 H_1:总体相关系数 $\rho>-0.570$,做单侧检验.

解 ☞

(1) $H_0:\rho=-0.570$, $H_1:\rho>-0.570$;

(2) $\alpha=0.05$;

(3) 由于是为了检验相关系数是否等于某个特定值,故采用 Z 分布(表 A.5)来寻找拒绝的临界值.查表 A.5 得 $z_{0.05}=1.645$,决策规则为:若 $z^*>1.645$,则拒绝 H_0;

(4) 利用表 A.10 将 $r=-0.410$ 和 $\rho=-0.570$ 转化成相应的 z 值,有 $z_r=-0.4356$, $z_\rho=-0.6475$.将这些 z 值及 $n=20$ 代入(19.36)式,得 z 的样本值

$$z^*=\frac{-0.4356-(-0.6475)}{\sqrt{1/17}}=0.8737$$

由表 A.5 知,对于标准正态分布,位于区间 $0\leqslant z\leqslant0.87$ 上方区域的面积为 0.3078,因而在 H_0 为真时,得到该 z 值的概率为

$$P(Z>0.87)=(0.5000-0.3078)=0.1922$$

(5) 由于 $z^*<1.645$,故接受 H_0.这一结论可由下述事实确认:在 $H_0:\rho=-0.570$ 为真时,得到 $r=-0.410$ 的概率大于 0.05.因而从总体相关系数 ρ 考虑,像是两个样本来自同一个总体.

19.36 重新考虑 19.24 中的关于雇员初始年薪与工作 10 年后的年薪之间的关系的研究,这次随机抽取容量为 16 的样本,计算得样本相关系数为 $r^*=0.400$.给定 $\alpha=0.01$,利用临界值决策规则,对零假设 H_0:总体相关系数 $\rho=0.264$,做双侧检验.

解 ☞

(1) $H_0:\rho=0.264$, $H_1:\rho\neq0.264$;

(2) $\alpha=0.01$;

(3) 由于是为了检验相关系数是否等于某个特定值,故采用 Z 分布(表 A.5)来寻找拒绝的临界值.查表 A.5 得 $z_{0.01/2}=2.575$,决策规则为:若 $z^*<-2.575$ 或 $z^*>2.575$,则拒绝 H_0;

(4) 利用表 A.10 将 $r=0.400$ 和 $\rho=0.264$ 转化成相应的 z 值,有 $z_r=0.4236$, $z_\rho=0.2704$.将这些 z 值及 $n=16$ 代入(19.36)式,得 z 的样本值

$$z^*=\frac{0.4236-0.2704}{\sqrt{1/13}}=0.5524$$

由表 A.5 知，对于标准正态分布，位于区间 $0 \leqslant z \leqslant 0.55$ 的上方的区域的面积为 0.2088，因而在 H_0 为真时，得到该 z 值的概率为

$$2[P(Z > 0.55)] = 2(0.5000 - 0.2088) = 2(0.2912) = 0.5824$$

(5) 由于 $-2.575 < z^* < 2.575$，故接受 H_0. 这一结论可由下述事实确认：在 $H_0: \rho = 0.264$ 为真时，得到 $r = 0.400$ 的概率大于 0.01. 因而从总体相关系数 ρ 考虑，像是两个样本来自同一个总体.

相关系数 r 的解释

19.37 通过说明鸟类的体温与呼吸速率之间的总变差中有多少是由这两个变量之间的线性关系而引起的，来解释习题 19.22 中的研究结果.

解 ☞ 决定系数 r^2 估计了两个变量间的总变差中，由于线性关系而引起的变差的大小. 这里 $r^* = 0.871$，$r^{*2} = 0.759$，因而我们说体温与呼吸速率之间的总变差中，有 75.9% 是由于两者间线性关系而引起的.

19.38 通过说明岛屿年龄与蜥蜴种类的数量之间的总变差中有多少是由这两个变量之间的线性关系而引起的，来解释习题 19.23 中的研究结果.

解 ☞ 在此研究中，$r^* = -0.570$，$r^{*2} = 0.325$. 因而我们说，岛屿年龄与蜥蜴种类数之间的总变差中有 32.5% 是由二者之间的线性关系而引起的.

19.39 通过说明两种年薪之间的总变差中有多少是两者之间的线性关系而引起的，来解释习题 19.24 中所描述的试验结果.

解 ☞ 在此研究中，$r^* = 0.264$，$r^{*2} \approx 0.070$. 因而我们说在两种年薪之间的总变差中，有 7.0% 是由两者之间的线性关系而引起的.

补充习题

最小二乘回归直线

19.40 一位生物学家想预测鸭子胚胎在各年龄段的身体活动水平. 她考察四个年龄段(按天计算)，每个年龄段观察三个胚胎，记录下每个胚胎在身体活动期间的平均持续时间(按秒计算)，结果如下(天：持续时间)：(5 天：1,3,1 秒)；(10 天：18,20,22 秒)；(15 天：32,36,37 秒)；(20 天：40,46,48 秒). 求这些数据的最小二乘回归直线？利用这条回归直线，预测 13 天的鸭子胚胎的身体活动持续时间？

答案： $\hat{\mu}^*_{Y|x} = -11.33 + 2.96x$，$\hat{\mu}_{Y|x=5} = 27.2$ 秒

19.41 某石油开采公司想建立最经济的探测钻井体系，因而需知道收益(单位：十亿美金)与格点距离(单位：英里)之间的关系，初步研究结果如下(格点距离：收益)：(2 英里：15.5,17.0,15.2)；(4 英里：11.0,13.1,9.8)；(6 英里：7.0,8.5,5.5)；(8 英里：2.3,1.8,1.7). 求这些数据的最小二乘回归直线？给定 5 英里的格点距离，预测该公司的收益？

答案： $\hat{\mu}^*_{Y|x} = 20.583 - 2.31x$，　$\hat{\mu}^*_{Y|x=5} = \$9.033$ 十亿

19.42 一位研究计算机商店的市场分析员，想根据商店每月广告费来预测该店在该月的营业收入. 她选取三种消费水准($200, $400, $600)，对于每种水平随机选取三家计算机商店，调查结果如下($广告费；$营业收入)：($200；$36,500)，($200；$35,400)，($200；$30,000)，($400；$41,200)，($400；$43,600)，($400；$46,000)，($600；$61,800)，($600；$52,500)，($600；$68,000). 求这些数据的最小二乘回归直线？利用此回归直线，来预测平均每月广告费为 $380 的计算机商店的平均月收入？

答案： $\hat{\mu}^*_{Y|x} = 19{,}311 + 67.0x$，　$\hat{\mu}^*_{Y|x=5} \approx \$44{,}771$

方差 $\sigma^2_{Y\cdot X}$ 的估计

19.43 考虑习题 19.40 中的鸭子胚胎试验，求估计的方差的最佳估计和该估计的标准误的最佳估计？

答案： $s^2_{Y\cdot x} = 7.20$ 秒^2，$s_{Y\cdot x} \approx 2.7$ 秒

19.44 考虑习题 19.41 中收益的初步研究，求估计量的方差的最佳估计及其标准误的最佳估计？

答案： $s^2_{Y\cdot x} = 1.252$ 十亿^2，$s_{Y\cdot x} = 1.12$ 十亿

19.45 考虑习题 19.42 中的广告费与营业收入之间关系的研究,求估计量的方差的最佳估计及其标准误的最佳估计?

答案:$s^2_{y\cdot x}$ = $ ^{2}26,553,571.4, $s_{y\cdot x}$ = $ 5,153.02

$\hat{a}$ 和 $\hat{b}$ 的方差

19.46 考虑习题 19.40 中的鸭子胚胎试验.求 $\hat{a}$ 和 $\hat{b}$ 的点估计的标准差?

答案:$s_{\hat{a}^*}$ = 1.90 秒, $s_{\hat{b}^*}$ = 0.139 秒/天

19.47 考虑习题 19.41 中的石油开采试验.求 $\hat{a}$ 和 $\hat{b}$ 的点估计的标准差?

答案:$s_{\hat{a}^*}$ = 0.791 十亿, $s_{\hat{b}^*}$ = 0.144 十亿/英里

19.48 考虑习题 19.42 中的广告费与营业收入试验.求 $\hat{a}$ 和 $\hat{b}$ 的点估计的标准差?

答案:$s_{\hat{a}^*}$ = $ 4,544.5, $s_{\hat{b}^*}$ = $ 10.52

a 和 b 的置信区间

19.49 考虑习题 19.40 中的鸭子胚胎试验.求 a、b 的 95% 置信区间?

答案:a:(−15.56 秒, −7.10 秒), b:(2.650 秒/天, 3.270 秒/天)

19.50 考虑习题 19.41 中的石油开采试验.求 a 和 b 的 90% 置信区间?

答案:a:($ 19.152 十亿, $ 22.016 十亿), b:(− $ 2.571 十亿/英里, − $ 2.049 十亿/英里)

19.51 考虑习题 19.42 中的广告费与营业收入试验.求 a 和 b 的 99% 置信区间?

答案:a:($ 3,409.8, $ 35,212.2), b:($ 30.19, $ 103.81)

方差 $\sigma^2_{Y\cdot X}$ 的置信区间

19.52 考虑习题 19.40 中的鸭子胚胎试验.求年龄与活动持续时间之间的线性关系的总体方差 $\sigma^2_{Y\cdot X}$ 的 95% 置信区间?

答案:(3.516 秒^2, 22.154 秒^2)

19.53 考虑习题 19.41 中的石油钻探试验,求收益与格点距离之间的线性关系的总体方差 $\sigma^2_{Y\cdot X}$ 的 90% 置信区间?

答案:($ ^{2}0.6838 十亿, $ ^{2}3.1777 十亿)

19.54 考虑习题 19.42 中的市场调研试验,求广告费与营业收入之间线性关系的总体方差 $\sigma^2_{Y\cdot X}$ 的 99% 置信区间?

答案:($ ^{2}9,165,433.9, $ ^{2}187,942,365.8)

Y 的期望值的预测区间

19.55 考虑习题 19.40 中的鸭子胚胎试验,现对 14 天的鸭子胚胎做试验,欲估计其身体活动平均持续时间,利用习题 19.40 给出的最小二乘回归直线做此预测,并利用 n = 12, $s_{y\cdot x}$ = 2.7 秒,求其 99% 预测区间?

答案:$\hat{\mu}_{Y\cdot x=14}$ = 30.11 秒,(21.180 秒,39.040 秒)

19.56 某石油钻探公司想采用格点距离为 3.0 英里的开采模式,并想预测其收益(单位:$. 试利用习题 19.41 给出的最小二乘回归直线做此预测,然后利用 n = 12 及 $s_{y\cdot x}$ = $ 1.12 十亿求其 95% 预测区间.

答案:$\hat{\mu}_{Y\cdot x=3}$ = $ 13.65 十亿,($ 10.974 十亿, $ 16.326 十亿)

19.57 一位计算机商店老板想知道:若他的平均每月广告费为 $ 150,他的平均月营业收入将是多少? 利用习题 19.42 给出最小二乘回归直线做此预测,并利用 n = 9, $s_{y\cdot x}$ = $ 5,153.02,求其 90% 预测区间?

答案:$\hat{\mu}_{Y\cdot x=150}$ = $ 29,361.0,($ 17,925.02, $ 40,796.98)

关于斜率 b 的假设检验

19.58 在习题 19.40 中,求得最小二乘回归直线的斜率 $\hat{b}$ 为 2.96 秒/天.在习题 19.46 中,求得 $\hat{b}$ 的标准差为 $s_{\hat{b}^*}$ = 0.139 秒/天.给定 α = 0.01,对假设 H_0:b = 0,和 H_1:b>0,做 t 检验,给出检验统计量的样本值,说明接受 H_0,还是拒绝 H_0,并给出在 H_0 为真时得到该 t^* 值的近似概率.

答案:t^* = 21.295,由于 t^* >2.764,故拒绝 H_0. P<0.0005

19.59 在习题 19.41 中,求得最小二乘回归直线的斜率 $\hat{b}$ 为 −2.31 十亿/英里.在习题 19.47 中,求得 $\hat{b}$ 的标准差为 $s_{\hat{b}^*}$ = $ 0.144 十亿/英里.给定 α = 0.05,对假设 H_0:b = 0,和 H_1:b<0,做 t 检验,给出检验统计量的样本值,说明接受 H_0,还是拒绝 H_0,并给出在 H_0 为真时得到该 t^* 值的近似概率.

答案：$t^* = -16.042$，由于 $t^* < -1.812$，故拒绝 H_0．$P<0.0005$

19.60　在习题 19.42 中，求得最小二乘回归直线的斜率 $\hat{b}$ 为 \$67.0．在习题 19.46 中，求得 $\hat{b}$ 的标准差为 $s_{\hat{b}} = \$10.52$．给定 $\alpha=0.05$，对假设 $H_0:b=0$，和 $H_1:b>0$，做 t 检验，给出检验统计量的样本值，说明接受 H_0，还是拒绝 H_0．并给出在 H_0 为真时得到该 t^* 值的近似概率．

答案：$t^*=6.369$，由于 $t^*>1.895$，故拒绝 H_0．$P<0.0005$

样本相关系数 r 的计算

19.61　我们想知道生长在沙漠中的木馏油灌木的生长密度是由什么决定的，研究假设之一是由降雨量决定．为此，随机选择 6 个研究地点，测量该地区的年降雨量（单位：厘米）及该地区的木馏油灌木的密度（单位：株/100 米^2），测量结果如下：（5.3 厘米；2 株/100 米^2），（15.8 厘米；3 株/100 米^2），（26.2 厘米；7 株/100 米^2），（29.9 厘米；5 株/100 米^2），（12.1 厘米；3 株/100 米^2），（18.0 厘米；4 株/100 米^2），求这两个变量的样本相关系数？

答案：$r^*=0.865$

19.62　一位城市公园规划者想在池塘中放养一些金鱼，为此她想知道金鱼需要多少阳光，由于含氧量是金鱼生存的一个重要因素，因而她想确定水温与含氧量之间的线性关系（假定存在线性关系），于是随机抽取 6 个池塘，测量每个池塘的水温（℃）和含氧量（毫升/1 升水），测量结果为：（4.5℃；9.2 毫升/1 升水），（9.3℃；8.1 毫升/1 升水），（36.6℃；5.8 毫升/1 升水），（24.0℃，5.0 毫升/1 升水），（15.8℃；7.2 毫升/1 升水），（30.1℃；5.1 毫升/1 升水）．求这两个变量之间的样本相关系数？

答案：$r^* = -0.883$

19.63　一位空气动力学的工程师检查模型飞机的设计，想知道飞机的体重与它在静止气体中的飞行速度的关系．他随机选取 6 架模型飞机，记录下每架飞机的重量（单位：千克）和最大飞行速度（单位：米/秒），结果如下：（0.035 千克；10.7 米/秒），（0.780 千克；9.9 米/秒），（0.093 千克；8.0 米/秒），（0.322 千克；13.0 米/秒），（0.275 千克，11.0 米/秒），（0.073 千克；13.5 米/秒）．求这两个变量之间的样本相关系数？

答案：$r^* = -0.123$

总体相关系数 ρ 的置信区间

19.64　考虑习题 19.61 中年降雨量与植物密度之间的关系，求总体相关系数 ρ 的 95% 置信区间？

答案：(0.179, 0.985)

19.65　考虑习题 19.62 中水温与含氧量之间的关系，求总体相关系数 ρ 的 90% 置信区间？

答案：(−0.981，−0.413)

19.66　考虑 19.63 中模型飞机的重量与速度之间的关系，求总体相关系数 ρ 的 95% 置信区间？

答案：(−0.850，−0.765)

相关系数 ρ 的假设检验

19.67　考虑习题 19.61 中年降雨量与植物密度的关系，已知 $r^*=0.865$．给定 $\alpha=0.05$，利用 r 值（表 A.11），对零假设 H_0：两变量间没有线性关系，做双侧检验，说明接受 H_0 还是拒绝 H_0，并给出得到该样本值的近似概率？

答案：由于 $r^*>0.811$，故拒绝 H_0，$0.02<P<0.05$

19.68　重新考虑习题 19.67 中的检验，但这次用 t 统计量．

答案：$t^*=3.448$，由于 $t^*>2.776$，故拒绝 H_0，$0.02<P<0.05$

19.69　考虑习题 19.62 中池塘水温与含氧量的关系，已知 $r^*=-0.883$．给定 $\alpha=0.05$，利用 r 值表（表 A.11），对零假设 H_0：两变量间无线性关系做双侧检验，说明接受 H_0 还是拒绝 H_0，并给出得到该样本值的近似概率．

答案：由于 $r^*<-0.811$，故拒绝 H_0，$0.01<P<0.02$

19.70　重新考虑习题 19.69 中的检验，但这次用 t 统计量．

答案：$t^*=-3.762$，由于 $t^*<-2.776$，故拒绝 H_0，$0.01<P<0.02$

19.71　考虑习题 19.63 中模型飞机体重与飞行速度的关系，已知 $r^*=-0.123$，给定 $\alpha=0.05$，利用 r 值表（表 A.11），对零假设 H_0：两变量间无线性关系，做双侧检验，说明接受 H_0 还是拒绝 H_0，并给出得到该样本值的近似概率．

答案：由于 $-0.811<r^*<0.811$，故接受 H_0，$P>0.10$

19.72　重新考虑习题19.71中的检验，但这次用t统计量.

答案：$t^* = -0.248$，由于$-2.776 < t^* < 2.776$，故接受H_0，$P > 0.20$

用Z分布检验假设$\rho = c$

19.73　继续考虑习题19.61中年降雨量与植物密度关系的研究，这次随机抽取容量为10的样本重做此试验，样本相关系数$r^* = 0.300$. 给定$\alpha = 0.05$，对零假设H_0：总体相关系数$\rho = 0.865$，做双侧检验，给出检验统计量的样本值，说明接受H_0还是拒绝H_0，并给出在H_0为真时，得到该z^*值的概率.

答案：$z^* = -2.655$，由于$z^* < -1.96$，故拒绝H_0，$P = 0.008$

19.74　继续考虑习题19.62中池塘的水温与含氧量之间关系的研究，这次以容量为12的样本重做此试验，样本相关系数$r^* = -0.640$. 给定$\alpha = 0.05$，对零假设H_0：总体相关系数$\rho = -0.883$，做双侧检验，给出检验统计量的样本值，说明接受H_0还是拒绝H_0，并给出在H_0为真时，得到该z^*值的概率.

答案：$z^* = 1.893$，由于$-1.96 < z^* < 1.96$，故接受H_0，$P = 0.06$

19.75　继续考虑习题19.63中模型飞机的体重与速度之间关系的研究，这次以容量为20的样本重做试验，样本相关系数$r^* = 0.248$，给定$\alpha = 0.05$，对零假设H_0：总体相关系数$\rho = -0.123$，做双侧检验，给出检验统计量的样本值，说明接受H_0还是拒绝H_0，并给出H_0为真时，得到该z^*值的概率.

答案：$z^* = 1.554$，由于$-1.96 < z^* < 1.96$，故接受H_0，$P = 0.12$

相关系数r的解释

19.76　通过说明在年降雨量和植物密度之间的总变差中有多少变差是由这两个变量之间的线性关系引起的，来解释习题19.61中的研究结果.

答案：74.8%

19.77　通过说明在池塘的水温与含氧量之间的总变差中有多少变差是由这两个变量之间的线性关系引起的，来解释习题19.62中的研究结果.

答案：78.0%

19.78　通过说明在模型飞机的体重与速度之间的总变差中有多少变差是由这两个变量之间的线性关系引起的，来解释习题19.63中的研究结果.

答案：1.5%

第二十章　非参数方法

20.1　非参数方法与参数方法

到目前为止,我们只讨论了用参数方法来检验关于总体参数(比如 μ 和 σ^2)的一些假设,且假定总体服从某一特定分布,比如正态分布.本章,我们将采用非参数方法来检验关于总体分布的,而不是关于总体参数的假设.一些非参数方法亦称为"分布自由"方法,即对总体分布不要求严格假定.

20.2　χ^2 检验

χ^2 检验适用于名义水平测量(见上册 2.4 节),样本是各类内的项目个数.比如,大学生可根据他们是一年级、二年级、三年级或四年级学生而分成四类,如果从某大学随机抽取一个样本并计算四类中每一类内的学生数,那么这些数据就是名义测量值,从而可采用 χ^2 检验.

χ^2 检验是用来评价各类内所观测到的频数是否显著不同于在一定假设下的期望频数.我们通常采用 χ^2 检验统计量:

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{20.1}$$

进行这种检验,这里 O_i 是第 i 类内观测到的项目个数,E_i 是原假设为真时,第 i 类内该项目的期望个数,$\sum$ 是对所有的 k 类数据求和.由于该统计量是基于差的平方,故样本值恒为正数.

当观测频数接近于期望频数时,检验统计量的抽样分布近似于自由度为 v 的单参数的 χ^2 分布(见 15.2 节),当自由度 $v<10$ 时,χ^2 分布向右偏斜,而当自由度增大时,χ^2 分布渐近于正态分布(见图 15-1).

对于 χ^2 检验,自由度由类别个数,而不是样本容量确定.然而,样本容量在确定 χ^2 分布如何渐近于检验统计量的抽样分布时,也是很重要的.样本容量越大,越接近于渐近分布.对于最小的样本容量,不同统计书有不同的建议,一个通常的要求是所有期望频数都大于等于 5.

χ^2 检验有下列假定:(1)独立抽样(见 13.4 节);(2)各类之间互不相容;(3)列出所有分类情形.这种检验适用于频数,而不适用于百分率或比例.

χ^2 检验的一般步骤是:由总体的随机样本得到观测值,由所假设的总体分布算得期望值,最后由 χ^2 统计量[即(20.1)式]算出 χ^2 值.观测频数与期望频数的差越大,χ^2 统计量值也越大,将该样本值与 χ^2 分布的值比较来确定一个 χ^2 统计量取该样本值的概率.表 A.7 给出了 χ^2 分布的临界值,当 $\chi^{2*} > \chi^2_{\alpha,v}$ 时,拒绝原假设,这里 χ^{2*} 为样本值,是由(20.1)式得到的,$\chi^2_{\alpha,v}$ 为给定显著性水平 α 和自由度 v 时的临界值.

下面,我们分别针对三类检验:**拟合优度、变量的独立性、比率的齐性**,给出 χ^2 检验的一般步骤.拟合优度检验是处理某一变量的两个或多个类内的样本数据,所检验的问题是:该变量的各类内的观测频数与在预先确定的理论模型下所预期的频数是否相同?变量的独立性检验处理的是两个变量,所检验的问题是:这两个变量是否相互独立?比率的齐次性检验处理的是来自于三个以上总体的二项比率,所检验的问题是:这些总体比率是否相同?

20.3　χ^2 拟合优度检验

χ^2 拟合优度检验是用来判断单个变量的 k 类内频数的分布是否与理论分布相同.术语

“拟合优度”意指观测(样本)频数“拟合”期望(理论)频数有多好,检验统计量为(20.1)式,自由度为 $v=k-1$,其中 k 为类的个数.

拟合优度检验的假设为

$$H_0:\text{总体分布}=\text{理论分布}$$

$$H_1:\text{总体分布}\neq\text{理论分布}$$

由样本得到的 χ^2 值的分布是离散分布,其中频数为整数(例如 1,4,20),而理论上 χ^2 分布是连续型分布.因而,χ^2 分析的结果只能是理论分布的近似.然而除了变量只有两类($v=k-1=2-1=1$)的情形之外,一般这种近似程度都是很好的.对于只有两个类别的检验,一般统计书中采用 **Yates 的连续性修正**

$$\chi^2=\sum_{i=1}^{k}\frac{(|O_i-E_i|-0.5)^2}{E_i} \tag{20.2}$$

这里每个观测值与期望值的偏差的绝对值又减少 0.5.

例 20.1 一位大学教育专家想知道,对于小型私立大学的四个年级:一年级、二年级、三年级和四年级,每个年级的学生数是否相等,是否构成均匀分布?为此从某校随机抽取 4,500 个本科生:一年级学生 1,200 人,二年级学生 1,100 人,三年级学生 1,150 人,四年级学生 1,050 人.给定显著性水平 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0:四个年级的学生频数服从均匀分布,做 χ^2 拟合检验.

解 ☞ (1) H_0:四个年级的学生频数构成均匀分布;

H_1:四个年级的学生频数不构成均匀分布;

(2) $\alpha=0.05$;

(3) 由表 A.7,给定 $\alpha=0.05$,$v=4-1=3$,得临界值 $\chi^2_{0.05,3}=7.81$.决策规则为:若 $\chi^{2*}>7.81$,则拒绝 H_0;

(4) 观测频数为:1,200, 1,100,1, 150,1,050.当原假设为真时,期望频数为:总样本容量除以总的类数,即(1/4)(4,500)=1,125.由(20.1)式得检验统计量的样本值为

$$\chi^{2*}=\frac{(1,200-1,125)^2}{1,125}+\frac{(1,100-1,125)^2}{1,125}+\frac{(1,150-1,125)^2}{1,125}+\frac{(1,050-1,125)^2}{1,125}$$

$$=11.111$$

由表 A.7 知,对于此分布得到该样本值的近似概率为 $0.010<P<0.025$.

(5) 由于 $\chi^{2*}>7.81$,故拒绝 H_0,接受 H_1.这一结果亦可由下一事实得以确认:当零假设为真时,得到 $\chi^{2*}=11.11$ 的概率小于 0.05.因此,四个年级的学生频数不构成均匀分布.

20.4 独立性的 χ^2 检验:列联表分析

列联表分析是用来判别两个名义变量是相互独立,还是互相关联的.一个变量的类按行排列,而另一变量的类按列排列,行列相交的地方称为格子,列联表的大小记为 $r\times c$,这里 r 表示行数,c 表示列数.例如:判断大学教育水平(按行排列的 4 类:一年级、二年级、三年级和四年级)是否会影响成绩(按列排列的 5 类:A,B,C,D,F),这是 4×5 列联表,要检验的假设为

$$H_0:\text{两个变量彼此独立}$$

$$H_1:\text{两个变量彼此不独立}$$

列联表给出了两个变量的类同时出现的观测频数与期望频数.观测频数来自样本观测数据,期望频数来自**边际频数**,即行总数与列总数.为求期望频数,我们有

$$E_{ij}=\frac{\text{第 } i \text{ 行总数}\times\text{第 } j \text{ 列总数}}{n} \tag{20.3}$$

这里 E_{ij} 是位于第 i 行第 j 列的格子的期望频数，n 为总的样本容量.

观测频数和期望频数放在同一个列联表中，每个格子内有两个值：观测频数和期望频数，期望频数一般用括弧括起.

当建立列联表后，可以用(20.1)式给出的 χ^2 检验统计量来比较观测频数与期望频数，即

$$X^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{20.4}$$

这里 O_{ij} 与 E_{ij} 分别表示位于第 i 行第 j 列的格子的观测频数与期望频数，$\sum$ 是对所有格子求和.该统计量近似服从自由度为 $v=(r-1)(k-1)$ 的 χ^2 分布，这里 r 为列联表的行数，k 为列联表的列数.比较(20.4)式给出的 χ^2 样本值与 χ^2 表(表 A.7)中的临界值 $\chi^2_{\alpha,v}$.由于我们仅对样本值是否大于期望值感兴趣，故当样本值(χ^{2*})大于临界值($\chi^2_{\alpha,v}$)时，拒绝原假设.

正如拟合优度检验一样，当 $v=1$ 时，在 2×2 的列联表中，χ^2 统计量往往过高估计两个变量之间的关系，然而我们可采用 Yates 的连续性修正而得到一个更保守的估计，其中每个格子的 O_i 与 E_i 的差的绝对值再减去 0.5[见(20.2)式].

例 20.2 一位大学教授想知道大学教育水平是否影响学生的哲学课成绩.上学期有 400 人修该门课：一年级 150 人，二年级 100 人，三年级 100 人，四年级 50 人.学期末，80 人得 A，148 人得 B，172 人得 C.成绩分布如表 20.1 所示，给定显著性水平 $\alpha=0.05$，利用临界值决策规则，对零假设 H_0：学生成绩与年级无关，做 χ^2 拟合检验.

表 20.1

	一年级	二年级	三年级	四年级	总数
A	25	12	25	18	80
B	47	31	50	20	148
C	78	57	25	12	172
总数	150	100	100	50	400

解 ☞ (1) H_0：学生成绩与年级无关，H_1：学生成绩与年级有关；

(2) $\alpha=0.05$；

(3) 由表 A.7，给定 $\alpha=0.05$，自由度 $v=(c-1)(r-1)=(4-1)(3-1)=6$，得临界值 $\chi^2_{0.05,6}=12.59$.决策规则为：当 $\chi^{2*}>12.59$ 时，拒绝 H_0；

(4) 对于此检验，我们采用列联表检验统计量[(20.4)式]，因而需对列联表的每个格子计算观测频数和期望频数.观测频数即为样本值(见表 20.1)，期望频数可由(20.3)式得到，例如：若年级对成绩无影响，则一年级学生得 A 的期望数为

$$\text{一年级学生得 A 的期望数} = \frac{(\text{一年级学生数})(\text{得 A 的学生总数})}{\text{学生样本总数}} = \frac{(150)(80)}{400} = 30$$

表 20.2 列出了观测频数和期望频数，每个格子包含两个值，上面的值为观测频数，下面用括弧括起的值为当两变量独立时的期望频数.由(20.4)式得

表 20.2

	一年级	二年级	三年级	四年级
A	25	12	25	18
	(30.0)	(20.0)	(20.0)	(10.0)
B	47	31	50	20
	(55.5)	(37.0)	(37.0)	(18.5)
C	78	57	25	12
	(64.5)	(43.0)	(43.0)	(21.5)

$$\chi^{2*}=\frac{(25-30.0)^2}{30.0}+\frac{(12-20.0)^2}{20.0}+\frac{(25-20.0)^2}{20.0}+\frac{(18-10.0)^2}{10.0}$$
$$+\frac{(47-55.5)^2}{55.5}+\frac{(31-37.0)^2}{37.0}+\frac{(50-37.0)^2}{37.0}+\frac{(20-18.5)^2}{18.5}$$
$$+\frac{(78-64.5)^2}{64.5}+\frac{(57-43.0)^2}{43.0}+\frac{(25-43)^2}{43}+\frac{(12-21.5)^2}{21.5}$$
$$=37.764$$

由表 A.7 知,对于此分布得到该样本值的近似概率为 $P<0.005$.

(5) 由于 $\chi^{2*}>12.59$,故拒绝 H_0,接受 H_1.这一结果亦可由下一事实确认:当原假设为真时,得到 $\chi^2=37.76$ 的概率小于 0.05.因此,学生哲学课成绩与年级有关.

20.5　k 个二项比率齐性的 χ^2 检验

关于独立性的 χ^2 检验(见 20.4 节)亦可用来检验 k 个总体的二项比率是否相同(即齐次的)[回忆 11.2 节,11.3 节,一个贝努里试验只有两个可能结果:成功或失败,二项比率是一系列贝努里试验中成功或失败出现的比率].在 17.26 节,对二个总体比率齐性,我们采用双侧检验

$$H_0: p_1=p_2$$
$$H_1: p_1\neq p_2$$

k 个比率齐性的 χ^2 检验将这个双样本检验推广到多个总体的情形,检验的假设为

$$H_0: p_1=p_2=p_3=\cdots=p_k$$
$$H_1: p_1, p_2, \cdots p_k \text{ 不全相等}$$

如果拒绝 H_0,则至少有一个总体的比率不同于其他总体的比率.

这个检验涉及到来自 k 个总体的 k 个独立样本,尽管零假设是关于比率的,样本数据却是呈频数形式,而不是比率.首先将频数在表内排成两行(成功和失败)k 列,然后采用列联表分析方法做检验.在原假设为真的条件下,利用(20.3)式计算每个格子的期望频数.由 χ^2 检验统计量[(20.4)式]比较期望频数与观测频数.最后,对于给定显著性水平 α 和自由度 $v=(r-1)(c-1)=(2-1)(k-1)=k-1$,比较样本值 χ^{2*} 与 χ^2 表(表 A.7)给出的临界值,当样本值大于临界值时,拒绝零假设.

例 20.3　一位研究人工养殖海水鱼的海洋生物学家,想知道雄性的比率是否随海水池而变化?他随机选择了四个海水池,计算每个海水池中鱼的雄性个数和雌性个数,结果如表 20.3 所示.给定 $\alpha=0.05$,对零假设 H_0:四个海水池中的鱼的雄性比例相同,做 χ^2 检验.

表 20.3

	池塘 1	池塘 2	池塘 3	池塘 4	总数
雄性	59	68	47	56	230
雌性	68	60	55	67	250
总数	127	128	102	123	480

解 ☞　(1) $H_0: p_1=p_2=p_3=p_4$, $H_1: p_1, p_2, p_3, p_4$ 不全相等;

(2) $\alpha=0.05$;

(3) 给定 $v=k-1=3$,由表 A.7 得 $\chi^2_{0.05,3}=7.81$,决策规则为:若 $\chi^{2*}>7.81$,则拒绝 H_0;

(4) 观测频数即为样本值,如表 20.3 所示,期望频数可由(20.3)式得到,两组频数列在表 20.4 中,其中期望频数列在观测频数的下方,以括弧括起.由(20.4)式可得

$$\chi^{2*}=\frac{(59-60.854)^2}{60.854}+\frac{(68-61.333)^2}{61.333}+\frac{(47-48.875)^2}{48.875}+\frac{(56-58.938)^2}{58.938}$$

$$+\frac{(68-66.146)^2}{66.146}+\frac{(60-66.667)^2}{66.667}+\frac{(55-53.125)^2}{53.125}+\frac{(67-64.062)^2}{64.062}$$

$$=1.919$$

由表 A.7 知，得到该样本值的概率为 $0.500<P<0.900$.

表 20.4

	池塘 1	池塘 2	池塘 3	池塘 4
雄性	59	68	47	56
	(60.854)	(61.333)	(48.875)	(58.938)
雌性	68	60	55	67
	(66.146)	(66.667)	(53.125)	(64.062)

(5) 由于 $\chi^{2\prime}<7.81$，故接受 H_0. 这一结果亦可由下述事实确认：当零假设为真时，得到 $\chi^2=1.92$ 的概率大于 0.05. 因此，四个海水池中鱼的雄性比例相同.

20.6　秩次检验

秩次检验对总体分布无更多假定，由于没有充分利用样本信息，且功效较低，因而只有当参数方法不适用时才采用秩次检验. 然而，若无任何假定，秩次检验仍无法运用. 下面各节所叙述的方法有两个假定：(1)抽样独立(见 13.4 节). 一些秩次检验假定各观测是独立的，而另外一些秩次检验则假定各对观测独立或观测的差独立；(2) 抽样总体有连续分布，即使样本是由离散型度量表述的. 当分布连续时不存在相等的观测，即观测值的秩不会出现相同的情形，或者检验观测值的差时，不会出现差为零的情形，而在实际抽样时，秩相等或差为零也会偶然发生的，对于每种方法将分别叙述如何处理这些相等情形，我们将在后面分别给出一些方法.

下面各节介绍的秩次检验适用于有序水平观测(见上册的 2.5 节)，可根据观测值具有某种特征的程度将其排秩. 序水平观测可以是关于大小的秩次信息(例如气味强度)；可以是按照某一标准关于方向大小的符号信息("+"或"−")；或者是涉及文字上的而不是大小的类的信息.

下面各节介绍的方法均基于秩. 对于小样本，样本观测值进行排秩可通过直接观测或利用计算器完成，而对于有许多等秩的大样本，则需通过计算机编程完成排秩.

20.7　单样本检验：Wilcoxon 符号秩检验

单样本的集中趋势的等级次序检验是用来确定总体的中位数是否等于某个假设值，即检验零假设 $H_0:\tilde{\mu}=\tilde{\mu}_0$. 这是基于中位数的定义：若总体的中位数等于假定值，则样本值中有一半大于该值，而另一半小于该值. 检验统计量是建立在**差的得分**基础之上的，差的得分可通过将每个观测值减去假定的中位数而得.

有两种单样本检验：符号检验和符号秩检验. 符号检验仅假定差的得分的总体分布是连续的，该检验是比较样本中正差得分(即观测值大于假设的中位数)个数与负差得分个数. 符号秩检验则假定差的得分的总体分布连续且关于中位数对称，它不仅考虑差的得分的符号，而且考虑了差的得分的大小(表现为秩). 这里，我们讨论的是 **Wilcoxon 单样本符号秩检验**，这是一种代替单样本 t 检验(见 16.19 节)的非参数方法. 对于大样本和小样本，Wilcoxon 符号秩检验有不同的检验程序.

对于小样本($n<30$)，Wilcoxon 符号秩检验程序如下：(1)随机抽取一个样本，对于每个样本观测值，减去零假设给定的中位数，将这些差的得分取绝对值，然后排序，记录下其所在位置作为其秩(从 1 到 n)，最小差的秩为 1，最大差的秩为 n. 若有两个差的得分相同，则取其所在位置的平均(如，若有两个差的得分排在第 4 位，则这两个差的得分的秩均为 $4\frac{1}{2}$，即 4 与 5 的

算术平均);(2)将差的得分的原始符号("+"或"−")附到其相应秩的前面,正秩的和记为 W_+,负秩的和记为 W_-. 若零假设为真,则每个秩应等可能地取"+"或"−",从而绝对值 $|W_+|$ 近似等于绝对值 $|W_-|$.

检验统计量可有如下几种形式:正秩的和(W_+),负秩的和(W_-),所有符号秩的和,或者两个秩和中的较小者,此处我们所采用的检验统计量 W 是 $|W_+|$ 与 $|W_-|$ 中较小的那一个值. 如果零假设不真,则 $\min\{|W_+|, |W_-|\}$ 将小于零假设为真时它的期望值. 检验统计量 W 的分布可以算出,附录中表 A.12(**Wilcoxon** W **的临界值**)给出了在各种显著性水平 α 和样本容量 n 下的临界值,当 W 的值小于临界值时,拒绝零假设.

例 20.4　一位电影评论员按照从 1(最差)到 10(最好)的标准评价电影,工作 20 年后,其评分中位数为 6.3. 他退休后,报纸主编雇用另一人继续从事其工作,工作一个月后,这位新评论员共观看了 10 部电影,其评分分别为:3.8,5.6,1.8,5.0,2.4,4.2,7.3,8.6,9.1,5.2. 该报主编想了解这位新评论员与前一位评论员在评分上是否有所不同. 给定显著性水平 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0:两位评论员的评分中位数相同,做双侧 Wilcoxon 符号秩检验.

解　(1) $H_0:\tilde{\mu}=6.3$, $H_1:\tilde{\mu}\neq 6.3$;

(2) $\alpha=0.05$;

(3) 对 $\alpha=0.05$, $n=10$ 的双侧 Wilcoxon 符号秩检验,查表 A.12 得临界值为 8. 决策规则为:若 $W<8$,则拒绝 H_0;

(4) W 的值可通过将每个观测值减去零假设给定的中位数(6.3),然后将这些差的绝对值排序、记录下其秩,再将差的"+"或"−"加到相应秩的前面而得,这些计算结果见表 20.5. 含有"+"的秩和 W_+ 为 16,含有"−"号的秩和 W_- 为 39. 由于 $16<39$,故检验统计量 $W=16$. 查表 A.12,得到 $W=16$ 的近似概率为 $P>0.10$;

表 20.5

观测值减去 $\tilde{\mu}_0$	差的绝对值	秩	符号秩
3.8−6.3=−2.5	2.5	7	−7
5.6−6.3=−0.7	0.7	1	−1
1.8−6.3=−4.5	4.5	10	−10
5.0−6.3=−1.3	1.3	4	−4
2.4−6.3=−3.9	3.9	9	−9
4.2−6.3=−2.1	2.1	5	−5
7.3−6.3=+1.0	1.0	2	+2
8.6−6.3=+2.3	2.3	6	+6
9.1−6.3=+2.8	2.8	8	+8
5.2−6.3=−1.1	1.1	3	−3
			$\|W_+\|=16$
			$\|W_-\|=39$

(5) 由于 $W>8$,故接受 H_0,这一结论可由下述事实得以确认:在零假设为真时,得到 $W=16$ 的概率大于 0.05,因而两位评论员的评分并无显著不同.

对于大样本($n\geqslant 30$),统计量 W 的分布近似于正态分布,用上面同样的方法得到 W 的值,将 W 的值(即 $\min\{|W_+|, |W_-|\}$)代入 Z 变换公式

$$Z=\frac{W-\dfrac{n(n+1)}{4}}{\sqrt{\dfrac{n(n+1)(2n+1)}{24}}} \tag{20.5}$$

由于 Wilcoxon 检验是研究两个秩和中较小者是否小于零假设时它的期望值，因此，决策规则为：对于单侧检验，当 $z^* < -z_\alpha$ 时，拒绝 H_0；对于双侧检验，当 $z^* < -z_{\alpha/2}$ 时，拒绝 H_0。这些临界值可由表 A.7 得到，这个表的用法参见 12.8 节。

例 20.5　一位历史教师给 200 人的班讲课，在过去几年中，所有学生成绩的中位数为 71.5。今年，他教 35 人的小班，想知道小班学生成绩的中位数是否高于大班学生成绩的中位数。35 个学生成绩的符号秩的和的绝对值为：$|W_+|=420$，$|W_-|=210$。给定 $\alpha=0.05$，对零假设 H_0：小班学生成绩的中位数与大班学生成绩的中位数相同，和备择假设 H_1：小班学生成绩的中位数高于大班学生成绩的中位数，利用临界值决策规则，做单侧 Wilcoxon 检验

解 ☞　(1) $H_0:\tilde{\mu}=71.5$，$\tilde{\mu}>71.5$；

(2) $\alpha=0.05$；

(3) 当 $n>30$ 时，通过 Z 变换将统计量 W 变换成相应的 z^*，再由表 A.5 得单侧检验的临界值 $z_{0.05}=1.645$。决策规则为：若 $z^*<-1.645$，则拒绝 H_0。

(4) 已知 $W=210$，其为两个秩和的绝对值中较小者，将 $W=210$，$n=35$ 代入(20.5)式，有

$$z^* = \frac{210-\dfrac{35(35+1)}{4}}{\sqrt{\dfrac{35(35+1)[2(35)+1]}{24}}} = -1.7198 \approx -1.72$$

为确定 P 值，可利用(16.5)式：$P=P[Z\leqslant(Z^*=-a)|H_0$ 为真]。由表 A.5，位于区间 $0\leqslant z\leqslant 1.72$ 上方区域的面积为 0.4573。由于原点右侧区域的面积为 0.5，故

$$P(Z>1.72)=0.5-0.4573=0.0427$$

由于 Z 分布具有对称性，从而有

$$P(Z<-1.72)=0.5-0.4573=0.0427\approx 0.043$$

(5) 由于 $z^*<-1.645$，故拒绝 H_0，接受 H_1。这一结果亦可由下述事实得以确认：P 值小于 0.05。因此，小班学生成绩好于大班学生成绩。

20.8　两样本检验：相依样本的 Wilcoxon 符号秩检验

两个相依样本的秩次检验是研究当观测值成对抽取时(见 17.1 节)，成对样本之间的差。这些检验与单样本检验几乎相同，不同之处在于：这里所考虑的差是两个样本中位数的差，而不是样本中位数与假设的中位数的差。零假设是：两个抽样总体的中位数之间无差异。**Wilcoxon成对样本符号秩检验**考虑到差的得分的符号与大小(秩)。当总体分布不能假定为正态分布，或者数据是有序水平而不是间隔水平或比率水平观测时(见上册 2.6，2.7 节)，这种非参数方法可代替两个成对样本 t 检验(见 17.12 节)。

Wilcoxon 检验步骤如下：随机抽取 n 对以同一尺度测量值的观测，计算出每对观测之间的差，将这些差值的绝对值按从 1(最小)到 n(最大)排序。若有两个差的得分相同，则取其平均秩。大于零的差的得分，其秩的前面付予"+"号，而小于零的差的得分，其秩的前面付予"－"号，若有差的得分为零(即这对观测之间无差异)，在计算过程中删去此对观测，然后对减小的样本容量做检验。求两个秩和：正的秩和(W_+)与负的秩和(W_-)。检验统计量为 $W=\min\{|W_+|,|W_-|\}$。依据小样本还是大样本，该检验统计量的分布有所不同。

对于小样本($n<30$)，表 A.12 给出对于各种显著性水平 α 和样本容量 n 下的 W 的临界值，然后将由样本观测值算出的 W 值与临界值比较，作出判断；对于大样本($n\geqslant30$)，将 W 的值变换为 Z 值，然后利用表 A.5 作判断(可参见 20.7 节的大样本情形的步骤)。

例 20.6　一位医学研究者想知道某项新型锻炼方式对 75 岁至 80 岁之间的妇女的脉搏速率是否有影响，随机抽取该年龄段的妇女 12 人，分别在 2 个月的锻炼以前与以后测量她们的脉搏速率，测量结果如下：(75,71)；(81,83)；(74,70)；(75,60)；(70,75)；(74,67)；(82,85)；(64,65)；(79,69)；(83,71)；(73,65)；(82,76)(第一个数值表示二个月前，第二个数值表示二个月后)。给定 $\alpha=0.05$，利用临界值决策规则，对零假

设 H_0:二个月前后脉搏中位数相同,和备择假设 H_1:二个月后脉搏中位数较低,做单侧 Wilcoxon 检验.

解 (1) $H_0: \tilde{\mu}_1 = \tilde{\mu}_2, H_1: \tilde{\mu}_1 > \tilde{\mu}_2$;

(2) $\alpha = 0.05$;

(3) 由于 $n<30$,用小样本的 W 统计量. 由表 A.12,对于 $n=12$ 的单侧检验,临界值为 $W_{0.05}=17$.决策规则为:若 $W<17$,则拒绝 H_0;

(4) 检验统计量 W 可通过对每个妇女二个月前的脉搏频率减去二个月后的脉搏频率得到. 将所有的差的绝对值排序,记录下相应的秩,并将每个差的原始符号加到这些秩的前面,结果见表 20.6. 由于 12 是两个秩和的绝对值中较小者,故 $W=12$,由表 A.12 知,得到 $W=12$ 的近似概率为 $0.010<P<0.025$.

表 20.6

脉搏频率(月前减月后)	差的绝对值	秩	符号秩
75−71=+4	4	4.5	+4.5
81−83=−2	2	2.0	−2.0
74−70=+4	4	4.5	+4.5
75−60=+15	15	12.0	+12.0
70−75=−5	5	6.0	−6.0
74−67=+7	7	8.0	+8.0
82−85=−3	3	3.0	−3.0
64−65=−1	1	1.0	−1.0
79−69=+10	10	10.0	+10.0
83−71=+12	12	11.0	+11.0
73−65=+8	8	9.0	+9.0
82−76=+6	6	7.0	+7.0
			$\lvert W_+ \rvert = 66.0$
			$\lvert W_- \rvert = 12.0$

(5) 由于 $W<17$,故拒绝 H_0,接受 H_1. 这一结果可由下述事实得以确认:当零假设为真时,得到 $W=12$ 的概率 $P<0.05$. 因此,二个月后脉搏中位数较低.

20.9 两样本检验:独立样本的 Mann-Whitney U 检验

Mann-Whitney U 检验或 **Wilcoxon-Mann-Whitney 检验**是一种秩次非参数检验,它主要是研究两个独立样本是否来自于同一总体,这是一种有序水平观测的秩和检验,假定两个样本独立且来自于有连续分布的总体. 零假设是两抽样总体的分布相同;备择假设是它们不同或者一个比另一个大(一个总体的个体数量大于另一个总体的个体数量). 当总体分布不能假定为正态分布,或者当观测数据是有序水平观测而不是间隔水平或比率水平观测时,对于独立样本(即不是成对样本),应该用 Mann-Whitney U 检验代替两样本的 t 检验.

检验步骤如下:(1) 将来自于样本 1 的 n_1 个得分(观测值)与来自于样本 2 的 n_2 个得分(观测值)合并,然后从小到大排序,记下其所在位置作为秩,若有一些得分相等,则取其所在位置的算术平均作为秩. 如果两个总体的分布相同,这两个样本的秩应当是随机混合的. 如果两个总体的分布不同,则一个样本的秩将高于另一个样本的秩. 记样本 1 的所有秩的和为 R_1,样本 2 的所有秩的和为 R_2,定义 U_1, U_2 为

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \tag{20.6}$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 \tag{20.7}$$

U_1, U_2 中的较小者,即为 **Mann-Whitney 检验统计量** U.

对于小样本($n_1 \leqslant 20$, $n_2 \leqslant 20$).附录中的表 A.13(**Mann-Whitney** U **的临界值**)给出了拒绝零假设的临界值(该表对于单侧检验和双侧检验,分别提供了两个显著水平下的临界值;对于其他的显著水平,可由别的表得到).对于给定水平 α,表中的行标 n_1 为两个样本容量中的较小者,列标 n_2 为两个样本容量中的较大者,行与列的交叉处,即为所需的临界值.如果零假设为真,则 U_1 应近似等于 U_2,若零假设不真,则 U_1, U_2 中的较小者(检验统计量 U),应小于在零假设为真时的期望值.

例 20.7　某位大学校长想知道高校文科教员与理科教员工作 10 年后是否有相同的薪金.为此,随机选取 10 位文科教员,调查结果如下:$35,000, $30,000, $45,000, $42,000, $50,000, $52,000, $25,000, $38,000, $40,500, $41,000;又随机选取 13 位理科教员,调查结果如下:$35,500, $40,000, $48,000, $53,000, $58,000, $57,000, $54,000, $49,000, $49,500, $51,000, $51,500, $57,500, $52,500.给定 $\alpha = 0.05$,利用临界值决策规则,对零假设 H_0:两种薪金的分布相同,做双侧 Mann-Whitney U 检验.

解　(1) H_0:文科教员与理科教员薪金的分布相同,H_1:文科教员与理科教员薪金的分布不同;

(2) $\alpha = 0.05$;

(3) 由于 $n_1 < 20$, $n_2 < 20$,故采用小样本的 Mann-Whitney U 检验.由表 A.13,对于 $n_1 = 10$, $n_2 = 13$, $\alpha = 0.05$ 的双侧检验,决策规则为:若 $U < 33$,则拒绝 H_0.

(4) 将两组样本观测值合并在一起进行排序,两组样本的秩和分别记为 R_1 和 R_2,如表 20.7 所示.

表 20.7

文科教员薪金	秩	理科教员薪金	秩
$35,000	3	$35,500	4
$30,000	2	$40,000	6
$45,000	10	$48,000	11
$42,000	9	$53,000	19
$50,000	14	$58,000	23
$52,000	17	$57,000	21
$25,000	1	$54,000	20
$38,000	5	$49,000	12
$40,500	7	$49,500	13
$41,000	8	$51,000	15
		$51,500	16
		$57,500	22
		$52,500	18
$n_1 = 10$	$R_1 = 76$	$n_2 = 13$	$R_2 = 200$

将 $n_1 = 10$, $n_2 = 13$, $R_1 = 76$ 代入(20.6)式得

$$U_1 = (10)(13) + \frac{10(10+1)}{2} - 76 = 109$$

将 $n_1 = 10$, $n_2 = 13$, $R_2 = 200$ 代入(20.7)式得

$$U_2 = (10)(13) + \frac{13(13+1)}{2} - 200 = 21$$

由于 $21 < 109$,故检验统计量 $U = 21$

(5) 由于 $U < 33$,故拒绝 H_0,接受 H_1.工作 10 年后,文科教员与理科教员的薪金的分布不同.

对于大样本($n_1 > 20$ 或 $n_2 > 20$),检验统计量 U 近似服从正态分布,从而可将 U 值变换

成相应的Z值

$$Z=\frac{U-\frac{n_1n_2}{2}}{\sqrt{\frac{(n_1(n_2)(n_1+n_2+1)}{12}}} \tag{20.8}$$

这里U是U_1与U_2中的较小者,n_1为样本1的容量,n_2为样本2的容量.

由于检验统计量U取两个样本值中的较小者,故决策规则为:对于左侧检验,若$z^*<-z_\alpha$,则拒绝H_0;对于双侧检验,若$z^*<-z_{\alpha/2}$,则拒绝H_0.

当两个样本间有许多等秩情形时(样本内存在等秩情形不影响结果),Z统计量应含有修正因子

$$\sum T_i=\sum\frac{t_i^3-t_i}{12} \tag{20.9}$$

这里t_i表示具有同一个给定秩的差的得分个数.例如,若有两个得分的秩为6,三个得分的秩为10,四个得分的秩为7,则t_i的值分别为2,3,4.从而有

$$Z=\frac{U-\frac{n_1n_2}{2}}{\sqrt{\left(\frac{n_1n_2}{(n_1+n_2)^2-(n_1+n_2)}\right)\left(\frac{(n_1+n_2)^3-(n_1+n_2)}{12}-\sum T\right)}} \tag{20.10}$$

修正因子使得Z的值略有增大,当检验统计量的样本值接近于临界值时,这是很重要的.

例20.8　一位城建规划者想知道郊区1的房屋是否比郊区2的房屋更贵,为此,他在郊区1随机抽取30个房屋,在郊区2随机抽取40个房屋,将这些样本观测值合并在一起排序(没有相同秩的情形),结果如下:$R_1=1,241$,$R_2=1,244$.给定$\alpha=0.05$,利用临界值决策规则,对零假设H_0:两郊区的房屋的价值相同,和备择假设H_1:郊区1的房屋价值高于郊区2的房屋价值,做单侧Mann-Whitney U检验.

解　(1) H_0:两郊区的房屋的价值相同,H_1:郊区1的房屋价值高于郊区2的房屋价值;

(2) $\alpha=0.05$;

(3) 由于$n_1>20$,$n_2>20$,故采用大样本的Mann-Whitney U检验,由表A.5,对此单侧检验有$z_{0.05}=1.645$,决策规则为:若$z^*<-1.645$,则拒绝H_0;

(4) 将$n_1=30$,$n_2=40$,$R_1=1241$,$R_2=1244$,分别代入(20.6),(20.7)式,得

$$U_1=(30)(40)+\frac{30(30+1)}{2}-1,241=424$$

及

$$U_2=(30)(40)+\frac{40(40+1)}{2}-1,244=776$$

由于424<776,故检验统计量U的值为424,因为不存在秩相等情形,不必采用修正因子,将$U=424$,$n_1=30$,$n_2=40$,代入(20.8)式.有

$$z^*=\frac{424-\frac{(30)(40)}{2}}{\sqrt{\frac{(30)(40)(30+40+1)}{12}}}=-2.089$$

为确定P值.可利用(16.5)式:$P=P[Z\leqslant(Z^*=-a)|H_0 \text{为真}]$.由表A.5,位于区间$0\leqslant z\leqslant 2.09$上方区域的面积为0.4817.由于原点右侧区域的面积为0.5,故

$$P[Z\geqslant(z^*=a)\quad H_0\text{为真}]=0.50-0.4817=0.0183$$

由于Z分布具有对称性,从而有

$$P[Z \leqslant (z^* = a) \mid H_0 \text{为真}] = 0.0183$$

(5) 由于 $z^* < -1.645$,故拒绝 H_0,接受 H_1.这一结果可由事实 $P<0.05$ 得以确认.因此,郊区1的房屋价值高于郊区2的房屋价值.

20.10　多样本检验:k 个独立样本的 Kruskal-Wallis H 检验

Kruskal-Wallis H 检验分析 k 个总体间的差别,这里 $k>2$.这种秩和检验是两样本的 Mann-Whitney U 检验的一种推广.正如方差分析中 F 检验那样(见18.12节),在 Kruskal-Wallis H 检验中,独立变量称为**因子**,而称因子的类为**处理**.例如,有氧运动对脉搏速率影响的研究中,有氧运动是独立变量或因子,有氧运动的各个水平是处理.零假设为:样本容量分别为 $n_1, n_2, \cdots, n_k$ 的 k 个独立处理来自于同一个总体,或者所有抽样总体有相同分布;备择假设是至少有一个总体分布不同于其他的总体分布.当总体分布不能假定为正态分布或者当数据是有序水平观测,而不是间隔水平或比率水平观测时,可采用这种非参数的秩次检验以代替单向方差分析 F 检验.

Kruskal-Wallis H 检验步骤:(1)将来自 k 个处理的所有观测值合并,然后将其从小到大排序,记录下它们所处的位置作为其秩(从1到 N,$N=\sum_{i=1}^{k} n_i$ 表示 k 个处理的观测值总数).若有一些秩相等,则取算术平均作为秩.对每种处理计算秩和,记为 R_i.(2)得出秩和后,寻找一种检验统计量.该统计量应刻画所得秩和与零假设为真时的期望值的差别.同第18章中的方差的参数分析法类似,这个检验统计量基于平方和,不同之处在于:在参数分析法中,我们用的是实际的样本观测值,而 Kruskal-Wallis 检验统计量 H 用的是秩的平方和.不做推导,直接给出检验统计量

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{k}\left(\frac{R_i^2}{n_i}\right) - 3(N+1) \tag{20.11}$$

这里 N 是样本观测总数,n_i 是第 i 个处理的样本观测值个数,R_i^2 是第 i 个处理的秩的平方和.

当得分中相等观测值的个数超过35%时,需采用修正因子对 H 作修正

$$1 - \frac{\sum (t_i^3 - t_i)}{N^3 - N} \tag{20.12}$$

这里对于某个秩 i,t_i 表示相等的观测值个数.将此修正因子代入检验统计量 H 有

$$H = \frac{\frac{12}{N(N+1)}\sum_{i=1}^{k}\left(\frac{R_i^2}{n_i}\right) - 3(N+1)}{1 - \frac{\sum (t_i^3 - t_i)}{N^3 - N}} \tag{20.13}$$

修正因子使得 H 的值略有增大,当 H 接近于检验的临界值时,这是很重要的.

H 的值标志着 k 个处理中秩的分布情况.H 值越大,秩的差别也越大.若 H 值大于在假定各处理来自同一总体时的期望值,则拒绝零假设.Kruskal-Wallis H 的分布依赖于处理的个数和每个处理内的样本观测值的个数.对于小样本($n\leqslant 5$),且只有三个处理时,H 分布的临界值可由精确概率表得到,即附录中的表A.14(**Kruskal-Wallis H 临界值表**).

例20.9　一位房地产商想知道城镇规模是否与住宅占地面积(以亩记)相关.为此,对于三种城镇规模:小型(不足10,000居民),中型(10,000至100,000居民),大型(100,000居民以上),每种规模随机抽取四户,住宅占地面积分别为:小型镇:5.0,0.8,4.0,4.5;中型镇:1.0,0.5,3.0,1.0;大型镇:0.25,0.30,0.50,1.40.给定 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0:三种类型城镇的住宅占地面积分布相同,做 Kruskal-Wallis H 检验.

解 (1) H_0:三种类型城镇的住宅占地面积分布相同,H_1:三种类型城镇的住宅占地面积分布不完全相同;

(2) $\alpha=0.05$;

(3) 由于 $k=3$,每组样本的观测数据个数都小于5,故采用小样本的 Kruskal-Wallis H 检验.由表A.14,对于 $n_1=n_2=n_3=4$,临界值为 $H_{0.05}=5.692$,决策规则为:若 $H>5.692$,则拒绝 H_0;

(4) 对于每个处理计算 R,从而求出检验统计量 H.首先将 k 个处理的所有观测值组合并按从小到大排序,记录下其秩,最小值的秩为1,最大值的秩为 N,这里 $N=n_1+n_2+n_3=4+4+4=12$ 表示所有处理的观测值总数,相同的观测值取其所在位置的平均作为秩,记 R_i 为第 i 个处理的秩和,计算结果见表20.8.由于相等的观测值不足35%(12个中有2个),故不必对相等的观测值的秩作修正.由(20.11)式得

$$H=\frac{12}{12(12+1)}\left(\frac{(38)^2}{4}+\frac{(25.5)^2}{4}+\frac{(14.5)^2}{4}\right)-3(12+1)=5.3173$$

由表A.14,得到该样本值的近似概率为 $0.05<P<0.10$;

表 20.8

小型城镇		中等城镇		大型城镇	
住宅占地面积(亩)	秩	住宅占地面积(亩)	秩	住宅占地面积(亩)	秩
5.00	12.0	1.00	6.5	0.25	1.0
0.80	5.0	0.50	3.5	0.30	2.0
4.00	10.0	3.00	9.0	0.50	3.5
4.50	11.0	1.00	6.5	1.40	8.0
$n_1=4$	$R_1=38$	$n_2=4$	$R_2=25.5$	$n_3=4$	$R_3=14.5$

(5) 由于 $H<5.692$,故接受 H_0.这一结果可由下述事实得以确认:当零假设为真时,得到 $H=5.3173$ 的概率大于0.05.因此,三种类型城镇的住宅占地面积分布相同.

对于大样本($n>5$)或者处理的个数多于三个,检验统计量 H 近似服从自由度为 $k-1$ 的 χ^2 分布(这里 k 为处理的个数).表A.7提供了 χ^2 分布的临界值.对于处理的个数与观测值的个数均很少时的检验,在一定的显著水平下,当 H 大于零假设为真时的预期值时,我们拒绝 H_0.

例 20.10 一位大学教务处长想知道年级(一年级、二年级、三年级和四年级)是否对学生在图书馆的时间有影响.为此,从每个年级随机抽取7名学生,并记下每个学生一周内在图书馆的时间.调查结果如下:一年级:11.5,(15.0,)12.0,6.0,7.0,9.0,10.5(小时/周);二年级:11.0,19.0,18.0,14.0,8.5,6.5,5.0(小时/周);三年级:12.5,8.5,4.0,13.0,19.5,20.0,10.0(小时/周);四年级:16.0,10.0,7.5,8.0,17.0,4.5,21.0(小时/周).给定 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0:四个年级的学生在图书馆的时间分布相同,做 Kruskal-Wallis 检验.

解 (1) H_0:四个年级的学生在图书馆的时间分布相同,H_1:四个年级的学生在图书馆的时间分布不完全相同;

(2) $\alpha=0.05$;

(3) 由于 $k>3,n>5$,利用 χ^2 分布求临界值.由表A.7,对于 $v=k-1=4-1=3$,$\chi^2_{0.05,3}=7.81$.决策规则为:当 $H>7.81$ 时,拒绝 H_0;

(4) 对于每个处理计算 R 从而求出检验统计量 H.首先将 k 个处理的所有观测值合并,再按从小到大排序,记录下其所在位置作为秩,最小值的秩为1,最大值的秩为 N,这里 $N=\sum_{i=1}^{4}n_i=7+7+7+7=28$ 表示所有处理的观测值总数,对于相等观测值取其所在位置的平均作为秩,记 R_i 为第 i 个处理的秩和,计算结果见表20.9.由于相等的观测值只有4个(占14.3%),故不必采用修正因子.由(20.11)式得

$$H=\frac{12}{28(28+1)}\left(\frac{(89.0)^2}{7}+\frac{(101.5)^2}{7}+\frac{(113.0)^2}{7}+\frac{(102.5)^2}{7}\right)-3(28+1)=0.6112$$

由表A.7,得到 $\chi^2=0.6112$ 的概率为 $0.500<P<0.900$;

表 20.9

一年级		二年级		三年级		四年级	
小时	秩	小时	秩	小时	秩	小时	秩
11.5	16.0	11.0	15.0	12.5	18.0	16.0	22.0
15.0	21.0	19.0	25.0	8.5	9.5	10.0	12.5
12.0	17.0	18.0	24.0	4.0	1.0	7.5	7.0
6.0	4.0	14.0	20.0	13.0	19.0	8.0	8.0
7.0	6.0	8.5	9.5	19.5	26.0	17.0	23.0
9.0	11.0	6.5	5.0	20.0	27.0	4.5	2.0
10.5	14.0	5.0	3.0	10.0	12.5	21.0	28.0
$n_1=7$	$R_1=89.0$	$n_2=7$	$R_2=101.5$	$n_3=7$	$R_3=113.0$	$n_4=7$	$R_4=102.5$

(5) 由于 $H<7.81$，故接受 H_0. 这一结果可由下述事实得以确认：当零假设为真时，得到 $H=0.6112$ 的概率大于 0.05. 因此，学生在图书馆的时间与年级无关.

20.11 Spearman 秩相关检验

当观测是序水平，而不是间隔水平或者比率水平时，**Spearman 秩相关检验**这种非参数方法可代替 Pearson 乘积矩相关检验(见 19.13 节)，它分析 n 个观测序对(x_i,y_i)的样本中，变量 X,Y 之间的关系. 对于每个变量，将其观测值排序计算秩，确定每对秩之间的相对应的程度. Pearson 检验是确定两变量线性相关的程度，而 Spearman 检验仅确定观测序对是否具有**单调增加的关系**，即当一个变量增大时，另一个变量是否也增大；或者具有**单调减少关系**，即当一个变量增大时，另一个变量减少.

Spearman 检验的零假设是两个变量之间没有关系. 而备择假设有三种可能：存在某种关系(双侧检验)，存在单调增的关系(单侧检验)，或者存在单调减的关系(单侧检验).

检验步骤如下：(1)对于每个变量，将观测值按从小到大排序，记下其秩，于是对于每个个体，相应于 X 有一个秩，相应于 Y 有一个秩，若某些秩相临但不相等，则取平均秩. 用这些成对的秩计算检验统计量，即 **Spearman 秩相关系数** r_S 或 **Spearman 秩次相关系数**，或 **Spearman *rho***. 当两变量间存在单调增的关系时，相关系数为正(+1.00 表示完全相关)；当两变量之间存在单调减关系时，相关系数为负(−1.00 表示完全相关)；当两变量之间没有关系时，相关系数为零. 正如 Pearson 相关系数一样，Spearman 秩相关系数的平方(r_S^2)可用来度量 X,Y 的变差中，由两变量间关系而引起的那部分变差所占比率. 依据是否存在等秩情形，有两种计算这种检验统计量的方法.

当某个变量的观测值有等秩情形时，可利用 Pearson 乘积矩相关系数公式[(19.31)式]来计算 Spearman 相关系数

$$r_S=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2(y-\bar{y})^2}}$$

其中对于每个个体，x 表一个变量的秩，y 表另一个变量的秩. 其计算式为[见(19.32)式]

$$r_S=\frac{\sum xy-\frac{\sum x\sum y}{n}}{\sqrt{\sum x^2-\frac{(\sum x)^2}{n}}\sqrt{\sum y^2-\frac{(\sum y)^2}{n}}}$$

若每个变量的观测值中都没有等秩情形出现,上述计算式可简化为

$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{20.14}$$

这里 n 表示秩对个数,d_i 表示每对秩之间的差,可由每个个体的两个秩相减得到.注意到每对秩越相近,d 值越小,从而上式的分数值接近于零,于是 r_S 接近于 +1;反之,若每对秩相似程度越弱,d 值越大,从而上式的分数由 1(不相关时 $r_S=0$)变到 −2,故 r_S 接近于 −1.

对于小样本($n<30$),r_S 的抽样分布的精确临界值可由附录中的表 A.15(**Spearman r_S 临界值表**)得到.由 n 确定表中的行位置,由 α 确定表中的列位置,行与列的交叉处即为所需的临界值.尽管表 A.15 给出的均为正的临界值,对于负相关,亦可由表 A.15 得到临界值,只不过差一个负号.

例 20.11 某保险公司的人事部经理想知道职员在大学里的表现是否是其工作表现的一个标志,为此,她随机抽取 12 名职员,并让他们的上司对他们的工作业绩进行评级,评定级别从 1(最好)到 12(最差),然后将他们所处级别与他们在大学里的成绩相比较,结果如表 20.10 所示(注意,对于工作表现,12 表示最差,而对于学校平均成绩,12 表示最好).给定 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0:工作表现与大学成绩无关,和 H_1:工作表现与大学成绩有单调减少的关系,做单侧 Spearmen 秩相关检验.

表 20.10

职员	秩:工作表现	学校平均成绩	秩:平均成绩	d_i	d_i^2
1	6	2.80	6	0	0
2	4	2.45	3	+1	1
3	2	3.55	11	−9	81
4	9	2.95	7	+2	4
5	12	2.50	4	+8	64
6	10	2.60	5	+5	25
7	5	3.15	9	−4	16
8	3	3.52	10	−7	49
9	1	3.75	12	−11	121
10	8	2.03	1	+7	49
11	7	3.00	8	−1	1
12	11	2.06	2	+9	81

解 (1) H_0:工作表现与大学成绩无关,H_1:工作表现与大学成绩有单调减少的关系;

(2) $\alpha=0.05$;

(3) 由于秩对个数小于 30,故采用表 A.15 得到临界值.对于 $\alpha=0.05$,$n=12$,决策规则为:若 $r_S<-0.503$,则拒绝 H_0;

(4) 没有等秩情形,故采用(20.14)式计算 Spearman 相关系数.首先将雇员的大学成绩从低(1)到高(12)排序,记录下其秩,将雇员的工作表现从最好(1)到最差(12)排序,记录下其秩.对于每个个体,将两个秩相减得 d_i,平方后得 $d_i^2(i=1,2,\cdots,12)$,计算结果见表 20.10.于是有

$$\sum_{i=1}^{12} d_i^2 = 492$$

将 $\sum_{1}^{12} d_i^2 = 492, n = 12$ 代入(20.14)式,有

$$r_S = 1 - \frac{6(492)}{12(12^2 - 1)} = -0.7203$$

查表 A.15 知,得到该样本值的近似概率为 $0.005 < P < 0.010$;

(5) 由于 $r_S < -0.503$,故拒绝 H_0,从而接受 H_1.这一结果可由下述事实得以确认:在 H_0 为真时,得到 $r_S = -0.7203$ 的概率小于 0.05.因而雇员的大学成绩与其工作表现负相关:大学成绩越好,工作表现也越好.两变量间的总变差中有 51.9% 是由两者负相关而引起的.

当样本容量较大时($n \geqslant 30$),r_S 近似服从标准差为 $\frac{1}{\sqrt{n-1}}$ 的正态分布,从而可将 r_S 值变换为 Z 值

$$z = \frac{r_S - 0}{\frac{1}{\sqrt{n-1}}} = r_S\sqrt{n-1} \tag{20.15}$$

这个 Z 统计量近似服从标准正态分布,临界值可由表 A.5 得到.

例 20.12 一位大学代数教师想知道学生的缺席次数与该学生的课堂表现是否负相关.在学期末,他将其班级的 35 名学生分别按缺席次数和课堂表现排序,缺席次数从最多到最少,课堂表现从最好到最差,并记录下他们的秩.记录结果见表 20.11.给定 $\alpha = 0.01$,利用临界值决策规则,对假设 H_0:课堂表现与缺席次数无关,和 H_1:课堂表现与缺席次数有单调减少的关系,做单侧 Spearman 秩相关检验.

解 ☞ (1) H_0:课堂表现与缺席次数无关,H_1:课堂表现与缺席次数有单调减少的关系;

(2) $\alpha = 0.01$;

(3) 由于秩对个数大于 30,故将 r_S 变换成相应的 Z 值,然后利用 Z 分布求临界值.由表 A.5 得,$z_{0.01} = 2.33$,决策规则为:若 $z^* < -2.33$,则拒绝 H_0;

(4) 由表 20.11 知,存在等秩情形,故采用 Pearson 乘积矩相关系数[即(19.31)式]来计算 Spearman 相关系数.所求的和列在表 20.11 中,将这些和代入(19.32)式有

$$r_S = \frac{10{,}229.5 - \frac{(630)(630)}{35}}{\sqrt{14{,}879 - \frac{630^2}{35}}\sqrt{14{,}910 - \frac{630^2}{35}}} = -0.3116$$

再利用(20.15)式将 r_S 变换为 z^* 得

$$z^* = -0.3116\sqrt{35-1} = -1.8169$$

为求出得到该样本值的概率,利用(16.5)式:$P = P[Z \leqslant (z^* = -a) \mid H_0 \text{ 为真}]$.由表 A.5 知,位于区间 $0 \leqslant z \leqslant 1.82$ 上方的区域面积为 0.4656.从而

$$P[Z \geqslant (z^* = a) \mid H_0 \text{ 为真}] = 0.5000 - 0.4656 = 0.0344$$

又因为 Z 分布关于原点对称,故得到 $z^* = -1.82$ 的概率为 $P = 0.0334$;

(5) 由于 $z^* > -2.33$,故接受 H_0,这一结果可由下述事实得以确认:在 H_0 为真时,得到该样本值的概率大于 0.01,因而学生的课堂表现与其缺席次数无关.此外,可认为课堂表现的变差中仅有 9.7% (-0.3116^2)是由缺席次数引起的.

表 20.11

学生	缺席		课堂表现		$(x_i)(y_i)$
	x_i	x_i^2	y_i	y_i^2	
1	10.5	110.25	25.0	625	262.5
2	17.5	306.25	11.0	121	192.5
3	28.0	784.00	17.0	289	476.0
4	24.0	576.00	18.0	324	432.0
5	9.0	81.00	33.0	1,089	297.0
6	33.5	1,122.25	35.0	1,225	1,172.5
7	2.0	4.00	1.0	1	2.0
8	19.0	361.00	5.0	25	95.0
9	17.5	306.25	22.0	484	385.0
10	1.0	1.00	34.0	1,156	34.0
11	35.0	1,225.00	3.0	9	105.0
12	7.5	56.25	27.0	729	202.5
13	25.0	625.00	12.0	144	300.0
14	10.5	100.25	4.0	16	42.0
15	33.5	1,122.25	10.0	100	335.0
16	3.0	9.00	26.0	676	78.0
17	22.0	484.00	19.0	361	418.0
18	27.0	729.00	13.0	169	351.0
19	7.5	56.25	6.0	36	45.0
20	29.0	841.00	29.0	841	841.0
21	15.5	240.25	28.0	784	434.0
22	20.0	400.00	16.0	256	320.0
23	4.5	20.25	32.0	1,024	144.0
24	12.0	144.00	24.0	576	288.0
25	31.0	961.00	2.0	4	62.0
26	13.0	169.00	23.0	529	299.0
27	32.0	1,024.00	9.0	81	288.0
28	6.0	36.00	30.0	900	180.0
29	30.0	900.00	21.0	441	630.0
30	21.0	441.00	15.0	225	315.0
31	14.0	196.00	7.0	49	98.0
32	4.5	20.25	31.0	961	139.5
33	26.0	676.00	20.0	400	520.0
34	15.5	240.25	8.0	64	124.0
35	23.0	529.00	14.0	196	322.0
$\sum$	630.0	14,897.00	630.0	14,910	11,229.5

习 题 解 答

χ^2 拟合优度检验

20.1 一位基因学家想知道某种紫菀花的颜色是否以孟德尔方式隐性遗传，如果这样，那么将两种粉色花杂交，后代将以白:粉:红＝1:2:1 的比例出现．该基因学家做杂交试验，植株了 100 株后代，结果发现：21 株白的，61 株粉的，18 株红的．给定 $\alpha=0.05$，利用临界

值决策规则，对零假设 H_0：花颜色分布比例为 1:2:1，做 χ^2 检验.

解 (1) H_0：总体分布比例为 1:2:1；H_1：总体分布比例不是 1:2:1；

(2) $\alpha = 0.05$；

(3) 给定 $\alpha = 0.05$，自由度 $v = k - 1 = 2$，查表 A.7，得临界值 $\chi^2_{0.05,2} = 5.99$，决策规则为：若 $\chi^{2*} > 5.99$，则拒绝 H_0；

(4) 利用(20.1)式计算检验统计量. 观测值分别为 21,61,18. 若比例为 1:2:1，则容量为 100 的样本中，期望值分别为$\frac{1}{4}(100)$，$\frac{1}{2}(100)$，$\frac{1}{4}(100) = 25,50,25$，于是

$$\chi^{2*} = \frac{(21-25)^2}{25} + \frac{(61-25)^2}{50} + \frac{(18-25)^2}{25} = 5.020$$

由表 A.7 知，得到该样本值的近似概率为 $0.050 < P < 0.100$；

(5) 由于 $\chi^{2*} < 5.99$，故接受 H_0，这一结论可由下述事实得以确认：在零假设为真时，得到 $\chi^{2*} = 5.02$的概率大于 0.05，因此，这种杂交试验的后代的分布比例为 1:2:1.

20.2 一位鱼类遗传学家想知道鱼唇形状的遗传（曲状为显性遗传、平直状为隐性遗传）与鳞的形状的遗传（平状为显性遗传，脊状为隐性遗传）是否由位于不同染色体上的两种不同基因控制. 如果结论成立，那么两种杂合体（每个均有鱼唇形状的两种基因与鱼鳞形状的两种基因）的后代将会出现 9:3:3:1 的比例. 她做杂交试验，并检查了 160 个后代，结果发现：72 条鱼有曲状唇和平状鳞，38 条鱼有曲状唇和脊状鳞，32 条鱼有平直状唇和平状鳞，18 条有平直状唇和脊状鳞，给定 $\alpha = 0.05$，利用临界值决策规则，对零假设 H_0：遗传特征的分布比例为 9:3:3:1，做 χ^2 检验.

解 (1) H_0：总体分布比例为 9:3:3:1；

H_1：总体分布比例不是 9:3:3:1；

(2) $\alpha = 0.05$；

(3) 给定 $\alpha = 0.05$，自由度 $v = k - 1 = 3$，查表 A.7，得临界值 $\chi^2_{0.05,3} = 7.81$，决策规则为：若 $\chi^{2*} > 7.81$，则拒绝 H_0；

(4) 利用(20.1)式计算检验统计量. 观测值分别为 72,38,32,18. 若比例为 9:3:3:1，则容量为 160 的样本中，期望值分别为$\frac{9}{16}(160)$，$\frac{3}{16}(160)$，$\frac{3}{16}(160)$，$\frac{1}{16}(160) = 90, 30, 30, 10$，于是

$$\chi^{2*} = \frac{(72-90)^2}{90} + \frac{(38-30)^2}{30} + \frac{(32-30)^2}{30} + \frac{(18-10)^2}{10} = 12.267$$

由表 A.7 知，得到该样本值的近似概率为 $0.005 < P < 0.010$；

(5) 由于 $\chi^{2*} > 7.81$，故拒绝 H_0，接受 H_1，这一结论可由下述事实得以确认：在零假设为真时，得到 $\chi^{2*} = 12.27$ 的概率小于 0.05，因此，这种杂交试验的后代的分布比例不是 9:3:3:1

20.3 一位历史学教授从周一至周五每天都给某个大班讲课，想知道每天的出席率是否相同，他检查了一周内每天学生出席人数，结果发现：星期一出席 283 人，星期二出席 332 人，星期三出席 360 人，星期四出席 307 人，星期五出席 243 人. 给定 $\alpha = 0.05$，利用临界值决策规则，对零假设 H_0：出席人数服从均匀分布（即每天的出席人数均相同），做 χ^2 检验.

解 (1) H_0：总体分布为均匀分布；

H_1：总体分布不是均匀分布；

(2) $\alpha = 0.05$；

(3) 给定 $\alpha = 0.05$，自由度 $v = k - 1 = 4$，查表 A.7，得临界值 $\chi^2_{0.05,4} = 9.49$，决策规则为：若 $\chi^{2*} > 9.49$，则拒绝 H_0；

(4) 利用(20.1)式计算检验统计量. 观测值为：283,332,360,307,243. 零假设为真时，（均匀分布的）期望值为平均出席人数：

$$E_i = \frac{283 + 332 + 360 + 307 + 243}{5} = 305$$

于是由(20.1)式得

$$\chi^{2*}=\frac{(283-305)^2}{305}+\frac{(332-305)^2}{305}+\frac{(360-305)^2}{305}+\frac{(307-305)^2}{305}+\frac{(243-305)^2}{305}=26.511$$

由表 A.7 知,得到该样本值的近似概率 $P<0.005$;

(5) 由于 $\chi^{2*}>9.49$,故拒绝 H_0,接受 H_1,这个结论可由下述事实得以确认:在 H_0 为真时,得到 $\chi^{2*}=26.511$ 的概率小于 0.05,故五天内的出席人数不服从均匀分布.

独立性的 χ^2 检验:列联表分析

20.4 某餐厅连锁店的经理想知道顾客满意程度是否与服务员的薪水有关,她随机抽取 100 名顾客,询问服务员的名字及服务状况是极佳、较好还是差,然后将服务员的薪水按低、中、高三个等级分类,调查结果如表 20.12 所示,无括弧的数值表示频数.给定 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0 服务质量与服务员的薪水无关,作 χ^2 检验.

解 (1) H_0:服务质量与服务员的薪水无关;

H_1:服务质量与服务员的薪水有关;

(2) $\alpha=0.05$;

(3) 给定 $\alpha=0.05$,自由度 $v=(c-1)(r-1)=(3-1)(3-1)=4$,这里 c 表示列数,r 表示行数,查表 A.7,有 $\chi^2_{0.05,4}=9.49$,决策规则为:若 $\chi^{2*}>9.49$,则拒绝 H_0;

(4) 检验统计量为(20.4)式,由(20.3)式得期望频数,即表 20.12 内的括弧内的数值,

表 20.12

	低等薪金	中等薪金	高等薪金	总数
服务极佳	9 (8.32)	10 (7.02)	7 (10.66)	26
服务较好	11 (16.32)	9 (13.77)	31 (20.91)	51
服务差	12 (7.36)	8 (6.21)	3 (9.43)	23
总数	32	27	41	100

将观测频数和期望频数代入(20.4)式,从而有

$$\chi^{2*}=\frac{(9-8.32)^2}{8.32}+\frac{(10-7.02)^2}{7.02}+\frac{(7-10.66)^2}{10.66}+\frac{(11-16.32)^2}{16.32}+\frac{(9-13.77)^2}{13.77}$$

$$+\frac{(31-20.91)^2}{20.91}+\frac{(12-7.36)^2}{7.36}+\frac{(8-6.21)^2}{6.21}+\frac{(3-9.43)^2}{9.43}$$

$$=18.658$$

由表 A.7,得到该样本值的近似概率为 $P<0.005$;

(5) 由于 $\chi^{2*}>9.49$,故拒绝 H_0,接受 H_1,这一结论可由下述事实得以确认:在零假设为真时,得到 $\chi^{2*}=18.66$ 的概率小于 0.05,因而顾客满意程度与服务员的薪水有关.

20.5 一位汽车制造商想知道顾客的年龄是否会影响其所购车的颜色,随机抽取 500 名购车者,记录下他们的年龄和所购车的颜色(蓝、红、白、黑),他将年龄分为三类:青年人(低于 30 岁),中年人(30 岁至 50 岁),老年人(50 岁以上),结果如表 20.13 中不在括号内的数值所示.给定 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0:所购车的颜色与顾客的年龄独立,做 χ^2 检验.

解 (1) H_0:所购车的颜色与顾客的年龄无关;

H_1:所购车的颜色与顾客的年龄有关;

(2) $\alpha=0.05$;

(3) 给定 $\alpha=0.05$,自由度 $v=(c-1)(r-1)=(4-1)(3-1)=6$,查表 A.7,有 $\chi^2_{0.05,6}=12.59$,

决策规则为：若 $\chi^{2*} > 12.59$，则拒绝 H_0；

(4) 检验统计量为(20.4)式，由(20.3)式得期望频数，如表 20.13 中的括号内数值所示，

表 20.13

	蓝	红	白	黑	总数
青年人	73	32	74	21	200
	(72.0)	(24.0)	(76.0)	(28.0)	
中年人	59	16	65	20	160
	(57.6)	19.2)	(60.8)	(22.4)	
老年人	48	12	51	29	140
	(50.4)	(16.8)	(53.2)	(19.6)	
总数	180	60	190	70	500

再由(20.4)式可得

$$\chi^{2*} = \frac{(73-72.0)^2}{72.0} + \frac{(32-24.0)^2}{24.0} + \frac{(74-76.0)^2}{76.0} + \frac{(21-28.0)^2}{28.0} + \frac{(59-57.6)^2}{57.6}$$

$$+ \frac{(16-19.2)^2}{19.2} + \frac{(65-60.8)^2}{60.8} + \frac{(20-22.4)^2}{22.4} + \frac{(48-50.4)^2}{50.4} + \frac{(12-16.8)^2}{16.8}$$

$$+ \frac{(51-53.2)^2}{53.2} + \frac{(29-19.6)^2}{19.6}$$

$$= 11.683$$

查表 A.7，得到该 χ^{2*} 值的概率为 $0.050 < P < 0.100$；

(5) 由于 $\chi^{2*} < 12.59$，故接受 H_0，这一结论可由下述事实得以确认：在零假设 H_0 为真时，得到 $\chi^{2*} = 11.68$ 的概率大于 0.05，因而，顾客的年龄与其所购车的颜色相互独立。

20.6 一位保险公司经理想知道车祸事故的频数是否与司机驾车的路程有关，她积累了过去10年内200名司机的信息资料，将10年内的车祸频数分成三类：小于5次，5次至10次之间，大于10次；将单程驾车距离分成三类：不足10英里，10至20英里之间，20英里以上，结果如表20.14中不在括号内的数值所示．给定 $\alpha = 0.01$，利用临界值决策规则，对零假设 H_0：车祸频数与驾车距离无关，作 χ^2 检验．

解 ☞　(1) H_0：车祸频数与驾车距离无关；

H_1：车祸频数与驾车距离有关；

(2) $\alpha = 0.01$；

(3) 给定 $\alpha = 0.01$，自由度 $v = (c-1)(r-1) = (3-1)(3-1) = 4$，查表 A.7，有 $\chi^2_{0.01,4} = 13.28$，决策规则为：若 $\chi^{2*} > 13.28$，则拒绝 H_0；

(4) 检验统计量为(20.4)式，由(20.3)式得期望频数，如表 20.14 中括号内数值所示，

表 20.14

	<5次	5–10次	>10次	总数
<10英里	42	31	30	103
	(47.895)	(31.415)	(23.690)	
10－20英里	32	23	5	60
	(27.900)	(18.300)	(13.800)	
>20英里	19	7	11	37
	(17.205)	(11.285)	(8.510)	
总数	93	61	46	200

于是有

$$\chi^{2*}=\frac{(42-47.895)^2}{47.895}+\frac{(31-31.415)^2}{31.415}+\frac{(30-23.690)^2}{23.690}+\frac{(32-27.900)^2}{27.900}+\frac{(23-18.300)^2}{18.300}$$

$$+\frac{(5-13.800)^2}{13.800}+\frac{(19-17.205)^2}{17.205}+\frac{(7-11.285)^2}{11.285}+\frac{(11-8.51)^2}{8.51}$$

$$=12.376$$

由表A.7,得到该样本值的概率为 $0.010<P<0.025$;

(5) 由于 $\chi^{2*}<13.28$,故接受 H_0,这一结论可由下述事实得以确认:在零假设为真时,得到 $\chi^{2*}=12.38$的概率大于0.01,故车祸频数与驾车距离相互独立.

k 个二项比率齐性的 χ^2 检验

20.7 一位10 000米竞赛的主办者想知道自4年前举办该赛事以来,女参赛者的比例是否有所提高,他查看了以往资料,结果发现:第一年,210名运动员中有98位女性;第二年,190名运动员中有100名女性;第三年,220名运动员中有133位女性;第四年,280名运动员中有169位女性.给定 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0:四年中女参赛者的比例没有发生变化,作 χ^2 检验.

解 (1) $H_0: p_1=p_2=p_3=p_4$,

$H_1: p_1, p_2, p_3, p_4$ 不全相等;

(2) $\alpha=0.05$;

(3) 给定 $\alpha=0.05$,自由度 $v=k-1=4-1=3$,查表A.7,得临界值 $\chi^2_{0.05,3}=7.81$,决策规则为:若 $\chi^{2*}>7.81$,则拒绝 H_0;

(4) 检验统计量为(20.4)式,利用边缘频数算出期望频数[(20.3)式],比如第一年比赛,女运动员的期望频数为

$$E_1=\frac{(\text{第一年比赛时运动员的总数})(\text{四年中女运动员总数})}{\text{四年中运动员总数}}$$

$$E_1=\frac{(210)(500)}{900}=116.667$$

表20.15列出了观测频数和期望频数.

表 20.15

	第一年	第二年	第三年	第四年	总数
女性	98	100	133	169	500
	(116.667)	(105.556)	(122.222)	(155.556)	
男性	112	90	87	111	400
	(93.333)	(84.444)	(97.778)	(124.444)	
总数	210	190	220	280	900

由(20.4)式得

$$\chi^{2*}=\frac{(98-116.667)^2}{116.667}+\frac{(100-105.556)^2}{105.556}+\frac{(133-122.222)^2}{122.222}+\frac{(169-155.556)^2}{155.556}$$

$$+\frac{(112-93.333)^2}{93.333}+\frac{(90-84.444)^2}{84.444}+\frac{(87-97.778)^2}{97.778}+\frac{(111-124.444)^2}{124.444}$$

$$=12.131$$

由表A.7,得到该样本值的概率为 $0.005<P<0.010$;

(5) 由于 $\chi^{2*}>7.81$,故拒绝 H_0,接受 H_1.这一结论可由下述事实得以确认:在零假设为真时,

得到 $\chi^{2*}=12.13$ 的概率小于0.05，因此女参赛者的比例在四年中有所变化.

20.8 一位棒球帽的营销者想知道，他在棒球比赛时的潜在市场是否随赛季的进展而变化. 他随机抽取100人，在五月、六月、七月、八月、九月他们进入棒球场地时，记录下他们是否戴棒球帽，结果发现戴棒球帽的人数（按月的顺序）分别为：59，61，65，68，47. 给定 $\alpha=0.01$，利用临界值决策规则，对零假设 H_0：戴棒球帽的人数比例不随季节中月份的变化而变化，作 χ^2 检验.

解 (1) $H_0: p_1=p_2=p_3=p_4=p_5$，

$H_1: p_1,p_2,p_3,p_4,p_5$ 不全相等；

(2) $\alpha=0.01$；

(3) 给定 $\alpha=0.01$，自由度 $v=k-1=5-1=4$，查表A.7，得临界值 $\chi^2_{0.01,4}=13.28$，决策规则为：若 $\chi^{2*}>13.28$，则拒绝 H_0；

(4) 检验统计量为(20.4)式，由边缘频数可得期望频数[见(20.3)式]，观测频数与期望频数如表20.16所示，

表 20.16

	五月	六月	七月	八月	九月	总数
戴棒球帽者	59	61	65	68	47	300
	(60)	(60)	(60)	(60)	(60)	
不戴棒球帽者	41	39	35	32	53	200
	(40)	(40)	(40)	(40)	(40)	
总数	100	100	100	100	100	500

由(20.4)式得

$$\chi^{2*}=\frac{(59-60)^2}{60}+\frac{(61-60)^2}{60}+\frac{(65-60)^2}{60}+\frac{(68-60)^2}{60}+\frac{(47-60)^2}{60}$$
$$+\frac{(41-40)^2}{40}+\frac{(39-40)^2}{40}+\frac{(35-40)^2}{40}+\frac{(32-40)^2}{40}+\frac{(53-40)^2}{40}$$
$$=10.833$$

查表A.7知，得到该样本值的概率为 $0.025<P<0.050$；

(5) 由于 $\chi^{2*}<13.28$，故在0.01的显著水平下接受 H_0，这一结论可由下述事实得以确认：在零假设 H_0 为真时，得到 $\chi^{2*}=10.833$ 的概率大于0.01，故在五个月期间，戴棒球帽人数比例没有变化.

20.9 某大学校管会想减少一年级学生的饮酒量，为了解一学年中学生的饮酒状况，他们在一年内的不同时间抽取了六个随机样本，每次对100名新生做调查，询问他们在上一周的饮酒量，答案分成两类：超过10次，10次或者更少，结果发现，被调查的100人中，在上周饮酒量超过10次的分布状况为：九月份40人，十月份31人，十一月份36人，二月份35人，五月份18人，四月份20人. 给定 $\alpha=0.01$，利用临界值决策规则，对零假设 H_0：新生在上周饮酒超过10次的比例在六次抽样中没有变化，作 χ^2 检验.

解 (1) $H_0: p_1=p_2=p_3=p_4=p_5=p_6$，

$H_1: p_1,p_2,p_3,p_4,p_5,p_6$ 不全相等；

(2) $\alpha=0.01$；

(3) 对于自由度 $v=k-1=6-1=5$，查表A.7，得临界值 $\chi^2_{0.01,5}=15.09$，决策规则为：若 $\chi^{2*}>15.09$，则拒绝 H_0；

(4) 检验统计量为(20.4)式，由边缘频数可得期望频数[见(20.3)式]，观测频数与期望频数如表20.17所示.

表 20.17

	九月	十月	十一月	一月	三月	四月	总数
超过 10 次	40	31	36	35	18	20	180
	(30)	(30)	(30)	(30)	(30)	(30)	
10 次或者更少	60	69	64	65	82	80	420
	(70)	(70)	(70)	(70)	(70)	(70)	
总数	100	100	100	100	100	100	600

由(20.4)式得

$$\chi^{2*}=\frac{(40-30)^2}{30}+\frac{(31-30)^2}{30}+\frac{(36-30)^2}{30}+\frac{(35-30)^2}{30}+\frac{(18-30)^2}{30}+\frac{(20-30)^2}{30}$$

$$+\frac{(60-70)^2}{70}+\frac{(69-70)^2}{70}+\frac{(64-70)^2}{70}+\frac{(65-70)^2}{70}+\frac{(82-70)^2}{70}+\frac{(80-70)^2}{70}$$

$$=19.333$$

查表 A.7 知,得到该样本值的概率为 $P<0.005$;

(5) 由于 $\chi^{2*}>15.09$,故拒绝 H_0,接受 H_1,这一结论可由下述事实得以确认:在零假设为真时,得到 $\chi^{2*}=19.33$ 的概率小于 0.01,因此,新生在上周饮酒超过 10 次的比例在六次抽样并非保持不变.

单样本检验:Wilcoxon 符号秩检验

20.10　一个由 12 位地质学家组成的探测队,对某种地质构成的露岩进行抽样,发现每 1,000 个化石标本中种类的中位数为 34. 最近发现一种新露岩,探测队想知道该露岩每 1,000个标本中的种类数是否与以前探测的露岩相同,于是每位地质学家从新露岩中收集样本并计算每1,000个样本中种类的中位数,结果如下:39,21,33,64,40,43,37,42,54,36,47,38. 给定 $\alpha=0.05$,利用临界值决策规则,对假设 $H_0:\tilde{\mu}=\tilde{\mu}_0$,和 $H_1:\tilde{\mu}\neq\tilde{\mu}_0$,做双侧 Wilcoxon 检验.

解 ☞　(1) $H_0:\tilde{\mu}=34$;

$H_1:\tilde{\mu}\neq34$;

(2) $\alpha=0.05$;

(3) 由于 $n<30$,采用 W 统计量的小样本检验. 由表 A. 12,对于 $n=12$ 的双侧检验,$W_{0.05}=14$,决策规则为:若 $W<14$,则拒绝 H_0;

(4) 将每个观测中位数减去零假设给定的中位数 34,对这些差取绝对值,并从小到大排序,记录下其秩,再将差的正号或负号加到其秩的前面,正秩的和记为 W_+,负秩的和记为 W_-,计算结果如表 20.18 所示. 由于 $10.5<67.5$,故 $W=10.5$,由表 A. 12 知,得到该样本值的近似概率为 $0.02<P<0.05$;

表 20.18

观测值减 $\tilde{\mu}_0$	差的绝对值	秩	符号秩
39−34=+5	5	5.0	+5.0
21−34=−13	13	9.5	−9.5
33−34=−1	1	1.0	−1.0
64−34=+30	30	12.0	+12.0
40−34=+6	6	6.0	+6.0
43−34=+9	9	8.0	+8.0
37−34=+3	3	3.0	+3.0
42−34=+8	8	7.0	+7.0

续表

观测值减 $\tilde{\mu}_0$	差的绝对值	秩	符号秩
$54-34=+20$	20	11.0	+11.0
$36-34=+2$	2	2.0	+2.0
$47-34=+13$	13	9.5	+9.5
$38-34=+4$	4	4.0	+4.0
			$\lvert W_+\rvert=67.5$
			$\lvert W_-\rvert=10.5$

(5) 由于 $W<14$，故拒绝 H_0，接受 H_1，这一结论可由下述事实得以确认：在零假设为真时，得到 $W=10.5$ 的概率小于0.05，因此在新露岩中每1,000个样本中种类的中位数不同于以前的露岩中种类中位数.

20.11 继续考虑习题20.10，地质学家们为判断新露岩每1,000个样本中是否比以前的露岩有更多的种类，决定从新露岩中抽取更多的样本，这次抽取50个样本，正的秩和为830，负的秩和为−445. 给定 $\alpha=0.05$，利用临界值决策规则，对假设 $H_0:\tilde{\mu}=\tilde{\mu}_0$ 和 $H_1:\tilde{\mu}>\tilde{\mu}_0$，做Wilcoxon检验.

解 ☞ (1) $H_0:\tilde{\mu}=34$，$H_1:\tilde{\mu}>34$；

(2) $\alpha=0.05$；

(3) 由于 $n>30$，将 W 统计量做 Z 变换，从而可由 Z 分布求得临界值. 由表A.5，对单侧检验，有 $z_{0.05}=1.645$，决策规则为：若 $z^*<-1.645$，则拒绝 H_0；

(4) 已知 $W=445$(两个和中绝对值较小者)，$n=50$，由(20.5)式得

$$z^*=\frac{445-\frac{50(50+1)}{4}}{\sqrt{\frac{50(50+1)[2(50)+1]}{24}}}=-1.858\approx-1.86$$

为确定 P 值，利用(16.5)式：$P=P[Z\leqslant(z^*=-a\mid H_0\text{ 为真})]$. 查表A.5知，位于区间 $0\leqslant z\leqslant 1.86$ 上方的区域面积为0.4686，由于原点右侧区域面积为0.5，故

$$P=P(Z\geqslant 1.86\mid H_0\text{ 为真})=0.5-0.4686=0.0314$$

由于 Z 分布关于原点对称，因此

$$P=P(Z\leqslant -1.86\mid H_0\text{ 为真})=0.5-0.4686=0.0314\approx 0.031$$

(5) 由于 $z^*<-1.645$，故拒绝 H_0，接受 H_1，这一结论可由下述事实得以确认；P 值小于0.05，因而与以前的露岩相比，新露岩每1,000样本种类的中位数较高.

20.12 A州中等收入家庭的税率为每 \$1,000 交纳 \$21.42，某家庭准备搬迁到B州，于是想知道那里的税率是否更高，该家庭在B州随机抽取12个县作调查，结果发现税率分别为：\$31.70，\$31.80，\$28.70，\$31.12，\$25.00，\$13.90，\$20.01，\$33.12，\$15.00，\$34.84，\$25.32，\$18.01(每 \$1,000). 给定 $\alpha=0.05$，利用临界值决策规则，对假设 $H_0:\tilde{\mu}=\tilde{\mu}_0$ 和 $H_1:\tilde{\mu}>\tilde{\mu}_0$，作单侧Wilcoxon检验.

解 ☞ (1) $H_0:\tilde{\mu}=\$21.42$，$H_1:\tilde{\mu}>\$21.42$；

(2) $\alpha=0.05$；

(3) 由于 $n<30$，故采用 W 统计量，对于单侧检验，$n=12$，查表A.13，得临界值 $W_{0.05}=17$，决策规则为：若 $W<17$，则拒绝 H_0；

(4) 将每个观测中位数减去零假设中位数(\$21.42)，将所有差的绝对值按从小到大排序，记下它们的秩，并将差的原始符号"+"号或"−"号加到该秩的前面，所有正秩的和记为 W_+；所有负秩的和记为 W_-；计算结果如表20.19所示，

表 20.19

观测值减 $\tilde{\mu}_0$	差的绝对值	秩	符号秩
\$31.70 − \$21.42 = + \$10.28	10.28	9	+9
\$31.80 − \$21.42 = + \$10.38	10.38	10	+10
\$28.70 − \$21.42 = + \$7.28	7.28	6	+6
\$31.12 − \$21.42 = + \$9.70	9.70	8	+8
\$25.00 − \$21.42 = + \$3.58	3.58	3	+3
\$13.90 − \$21.42 = − \$7.52	7.52	7	−7
\$20.01 − \$21.42 = − \$1.41	1.41	1	−1
\$33.12 − \$21.42 = + \$11.70	11.70	11	+11
\$15.00 − \$21.42 = − \$6.42	6.42	5	−5
\$34.84 − \$21.42 = + \$13.42	13.42	12	+12
\$25.32 − \$21.42 = + \$3.90	3.90	4	+4
\$18.01 − \$21.42 = − \$3.41	3.41	2	−2
			$\|W_+\| = 63$
			$\|W_-\| = 15$

由于 15<63,故 $W = 15$,查表 A.13 知,得到该样本值的概率为 $0.025 < P < 0.050$;

(5) 由于 $W < 17$,故拒绝 H_0,接受 H_1,这一结论可由下述事实得以确认:在零假设为真时,得到 $W = 15$ 的概率小于 0.05,故 B 州的税率高于 A 州的税率.

20.13 一位医学研究者想知道空调是否会增加人的能量消耗,已知某大城市中,在暖热的环境下,每位女性每天消耗的能量的中位数为 2,500 千卡,她随机选取 40 位妇女,让她们在有空调的凉快的公寓里生活 10 天,然后在第 11 天,记录下她们的食物消耗量,计算出观测值与中位数 2,500 的差,并将之排序,结果发现:$|W_+| = 520$, $|W_-| = 300$. 给定 $\alpha = 0.05$,利用临界值决策规则,对假设 $H_0: \tilde{\mu} = \tilde{\mu}_0$ 和 $H_1: \tilde{\mu} > \tilde{\mu}_0$,做单侧 Wilcoxon 检验.

解 (1) $H_0: \tilde{\mu} = 2500$, $H_1: \tilde{\mu} > 2500$;

(2) $\alpha = 0.05$;

(3) 由于 $n > 30$,用 Z 变换形式的 W 统计量.由表 A.5,对于 $\alpha = 0.05$ 的单侧检验,$z_{0.05} = 1.645$,决策规则为:若 $z^* < -1.645$,则拒绝 H_0;

(4) 由于 300<520,故 $W = 300$,将 $W = 300$, $n = 40$ 代入(20.5)式得

$$z^* = \frac{300 - \dfrac{40(40+1)}{4}}{\sqrt{\dfrac{40(40+1)[2(40)+1]}{24}}} = -1.479 \approx -1.48$$

为确定 P 值,利用(16.5)式:$P = P[Z \leqslant (z^* = -a) | H_0 \text{为真}]$.由表 A.5 知,位于区间 $0 \leqslant z \leqslant 1.48$ 上方的区域面积为 0.4306,由于原点右侧区域的面积为 0.5,故

$$P = P(Z \geqslant 1.48 | H_0 \text{为真}) = 0.5 - 0.4306 = 0.0694$$

而 Z 分布关于原点对称,故

$$P = P(Z \leqslant -1.48 | H_0 \text{为真}) = 0.5 - 0.4306 = 0.0694 \approx 0.069$$

(5) 由于 $z^* > -1.645$,故接受 H_0,这一结论可以下述事实得以确认:P 值大于 0.05,故空调装置并未增加妇女对能量的消耗.

两样本检验:相依样本的 Wilcoxon 符号秩检验

20.14 考虑例 20.6 中的运动试验,在 75 至 80 岁的妇女中重新抽取容量为 35 的样本,经计算得:正秩的和为 402,负秩的和为 −228. 给定 $\alpha = 0.01$,利用临界值决策规则,对假设 H_0:锻炼前后脉搏速率中位数相同,和 H_1:锻炼后脉搏速率中位数($\tilde{\mu}_2$)小于锻炼前脉搏速率中位数($\tilde{\mu}_1$).

解 ☞　(1) $H_0:\tilde{\mu}_1=\tilde{\mu}_2, H_1:\tilde{\mu}_1>\tilde{\mu}_2$；

(2) $\alpha=0.01$；

(3) 由于 $n>30$，将 W 统计量做 Z 变换，利用 Z 分布求临界值. 由表 A.5 知，对于单侧检验，$z_{0.01}=2.33$，决策规则为：若 $z^*<-2.33$，则拒绝 H_0；

(4) 由于 228<402，故 $W=228$. 将 $W=228, n=35$ 代入(20.5)式，得

$$z^*=\frac{228-\dfrac{35(35+1)}{4}}{\sqrt{\dfrac{35(35+1)[2(35)+1]}{24}}}=-1.425$$

为确定 P 值，利用(16.5)式：$P=P[Z\leqslant(z^*=-a)\mid H_0\text{ 为真}]$. 由表 A.5 知，位于区间 $0\leqslant z\leqslant 1.43$ 上方区域的面积为 0.4236，由于原点右侧区域的面积为 0.5，故

$$P=P(Z\geqslant 1.43\mid H_0\text{ 为真})=0.5-0.4236=0.0764$$

而 Z 分布关于原点对称，故

$$P=P(Z\leqslant -1.43\mid H_0\text{ 为真})=0.5-0.4236=0.0764\approx 0.076$$

(5) 由于 $z^*>-2.33$，故在 0.01 的显著水平下接受 H_0，这一结论可由下述事实得以确认：得到该样本值 -1.425 的概率大于 0.01，因而，锻炼后脉搏速率并不低于锻炼前脉搏速率.

20.15　一位牛奶场主想知道某种新荷尔蒙激素是否会增加奶牛的产奶量，为此，他随机抽取 10 头奶牛做观测，对于每头奶牛，记录其产奶量，经这种新荷尔蒙激素处理一周后，再记下其产奶量，测量结果如下：(30, 34)；(25, 35)；(22, 27)；(25, 24)；(23, 25)；(34, 26)；(33, 24)；(30, 24)；(24, 27)；(32, 21)，每对数值中的"第一个"表示试验前的测量值，"第二个"表示试验后的测量值. 给定 $\alpha=0.05$，利用临界值决策规则，对假设 H_0：荷尔蒙激素处理前后牛奶产量不变，和 H_1：荷尔蒙激素处理后牛奶产量高于处理前牛奶产量，做单侧 Wilcoxon 检验.

解 ☞　(1) $H_0:\tilde{\mu}_1=\tilde{\mu}_2\quad H_1:\tilde{\mu}_1<\tilde{\mu}_2$；

(2) $\alpha=0.05$；

(3) 由于 $n<30$，利用小样本的 W 统计量. 由表 A.12，对于 $n=10$ 的单侧检验. 临界值 $W_{0.05}=11$，决策规则为：若 $W<11$，则拒绝 H_0；

(4) 将每头奶牛试验前的牛奶产量减去试验后的牛奶产量，并取绝对值，再将这些绝对值排序，记下其秩，并将差的原始符号"+"或"−"号加到其秩前面，正秩的和记为 W_+，负秩的和记为 W_-，计算结果如表 20.20 所示.

表 20.20

牛奶：试验前减试验后	差的绝对值	秩	符号秩
30 − 34 = −4	4	4	−4
25 − 35 = −10	10	9	−9
22 − 27 = −5	5	5	−5
25 − 24 = +1	1	1	+1
23 − 25 = −2	2	2	−2
34 − 26 = +8	8	7	+7
33 − 24 = +9	9	8	+8
30 − 24 = +6	6	6	+6
24 − 27 = −3	3	3	−3
32 − 21 = +11	11	10	+10
			$\lvert W_+\rvert=32$
			$\lvert W_-\rvert=23$

由于 23<32,故 $W=23$,查表 A.12 知,得到该样本值的概率为 $P>0.05$;

(5) 由于 $W>11$,故接受 H_0,这一结论可由下述事实得以确认:在零假设为真时,得到 $W=32$ 的概率大于 0.05,故荷尔蒙激素处理前后牛奶产量相同.

20.16 一位制鞋厂主想知道他的新型靴子是否比目前正在销售的靴子更耐磨,为此,他随机抽取 13 位徒步旅行者,每人随机选一只脚穿新型靴子,另一只脚穿正在销售的靴子,记录下每只靴子磨损到某种程度时所走过的路程(以英里为单位),观测结果如下:(460,530);(420,525);(520,500);(515,505);(490,520);(490,450);(500,495);(550,575);(480,474);(530,515);(518,490);(515,480);(475,493),每对数据中,第一个为目前正在销售的靴子(靴子 1)的数据,第二个为新型靴子(靴子 2)的数据.给定 $\alpha=0.05$,利用临界值决策规则,对假设 H_0:两种靴子耐磨时间的中位数相同,和 H_1:新型靴子耐磨时间的中位数较大,作 Wilcoxon 检验.

解 (1) $H_0:\tilde{\mu}_1=\tilde{\mu}_2$, $H_1:\tilde{\mu}_1<\tilde{\mu}_2$;

(2) $\alpha=0.05$;

(3) 由于 $n<30$,采用小样本的 W 统计量.由表 A.12,对于 $n=13$ 的单侧检验,临界值 $W_{0.05}=21$,决策规则为:若 $W<21$,则拒绝 H_0;

(4) 将每对数据中第一个数值减去第二个数值,对所有差取绝对值后按从小到大顺序排序,记下其差的绝对值的秩,并将差的原始符号"+"或"-"号加到其秩前面,正秩的和记为 W_+,负秩的和记为 W_-,计算结果见表 20.21.由于 45<46,故 $W=45$,由表 A.12 知,得到该样本值的概率为 $P>0.05$;

表 20.21

英里:靴子 1 减靴子 2	差的绝对值	秩	符号秩
460 - 530 = -70	70	12	-12
420 - 525 = -105	105	13	-13
520 - 500 = +20	20	6	+6
515 - 505 = +10	10	3	+3
490 - 520 = -30	30	9	-9
490 - 450 = +40	40	11	+11
500 - 495 = +5	5	1	+1
550 - 575 = -25	25	7	-7
480 - 474 = +6	6	2	+2
530 - 515 = +15	15	4	+4
518 - 490 = +28	28	8	+8
515 - 480 = +35	35	10	+10
475 - 493 = -18	18	5	-5
			$\|W_+\|=45$
			$\|W_-\|=46$

(5) 由于 $W>21$,故接受 H_0,这一结论可由下述事实得以确认:在零假设为真时,得到 $W=45$ 的概率大于 0.05,因此,新型靴并不比旧型靴耐磨.

20.17 再考虑习题 20.16,该制造商重新做此试验,随机抽取 50 位徒步旅行者,将目前正在销售的靴子(靴子 1)走过的路程减去新型靴子(靴子 2)走过的路程,对所有差取绝对值后按从小到大排序,记下其秩,并将差的原始符号"+"或"-"号加到其秩前,从而得到下列符号秩和:$W_+=850$, $W_-=-425$.给定 $\alpha=0.05$,利用临界值决策规则,对假设 H_0:两种靴子耐磨时间的中位数相同,和 H_1:新型靴子耐磨时间的中位数较大,做 Wilcoxon 检验.

解 ☞ (1) $H_0:\tilde{\mu}_1=\tilde{\mu}_2, H_1:\tilde{\mu}_1<\tilde{\mu}_2$;

(2) $\alpha=0.05$;

(3) 由于 $n>30$,故将 W 统计量做 Z 变换,然后利用 Z 分布求临界值.由表 A.5,对于单侧检验,临界值 $z_{0.05}=1.645$,决策规则为:若 $z^*<-1.645$,则拒绝 H_0;

(4) 由于 425<850,故 $W=425$,将 $W=425$, $n=50$ 代入(20.5)式得

$$z^*=\frac{425-\frac{50(50+1)}{4}}{\sqrt{\frac{50(50+1)[2(50)+1]}{24}}}=-2.051\approx-2.05$$

为确定 P 值,利用(16.5)式:$P=P[Z\leqslant(z^*=-a)|H_0$ 为真].查表 A.5 知,位于区间 $0\leqslant z\leqslant 2.05$ 上方的区域面积为 0.4798,由于原点右侧区域面积为 0.5,故

$$P=P(Z>2.05)=0.5-0.4798=0.0202$$

利用 Z 分布的对称性,有

$$P=P(Z<-2.05)=0.5-0.4798=0.0202$$

(5) 由于 $z^*<-1.645$,故拒绝 H_0,接受 H_1.这一结论可由下述事实得以确认:在零假设为真时,得到 $z^*=-2.05$ 的概率小于 0.05,因而,新型靴子比正在销售的靴子更耐磨.

两样本检验:独立样本的 Mann-Whitney U 检验

20.18 一位运动学家想知道某两种品牌的运动饮料对于参加 10 公里精英赛的运动员的效果,她随机抽取 8 名女运动员喝 A 饮料,另外 8 名女运动员喝 B 饮料,然后这 16 名女运动员参加 10 公里赛跑,结果如下(单位:分钟):喝 A 饮料的成绩分别为:33.30,30.10,38.62,38.94,42.63,41.96,46.30,43.25;喝 B 饮料的成绩分别为:31.62,46.33,31.82,40.21,45.72,39.80,45.60,41.25.给定 $\alpha=0.05$,利用临界值决策规则,对零假设 H_0:两组赛跑时间总体分布相同,作双侧 Mann-Whitney U 检验.

解 ☞ (1) H_0:用 A 饮料的女运动员的赛跑时间分布与用 B 饮料的女运动员的赛跑时间分布相同,H_1:用 A 饮料的女运动员的赛跑时间分布与用 B 饮料的女运动员赛跑时间分布不同;

(2) $\alpha=0.05$;

(3) 由于 $n_1<20$, $n_2<20$,故采用小样本的 Mann-Whitney U 检验.由表 A.13,对于 $n_1=8$, $n_2=8$的双侧检验,临界值 $U_{0.05}=13$.决策规则为:若 $U<13$,则拒绝 H_0;

(4) 我们利用(20.6)与(20.7)式计算 U_1, U_2,已知 $n_1=8$, $n_2=8$,为求 R_1, R_2,将两组样本值合并,然后记录下这些值的秩.记两组样本值的秩的和分别为 R_1, R_2,计算结果如表 20.22 所示.将

表 20.22

时间:A 饮料	秩	时间:B 饮料	秩
33.30	4	31.62	2
30.10	1	46.33	16
38.62	5	31.82	3
38.94	6	40.21	8
42.63	11	45.72	14
41.96	10	39.80	7
46.30	15	45.60	13
43.25	12	41.25	9
$n_1=8$	$R_1=64$	$n_2=8$	$R_2=72$

$n_1=8$, $n_2=8$, $R_1=64$ 代入(20.6)式得

$$U_1=(8)(8)+\frac{(8)(8+1)}{2}-64=36$$

将 $n_1=8, n_2=8, R_2=72$ 代入(20.7)式有

$$U_2=(8)(8)+\frac{(8)(8+1)}{2}-72=28$$

由于 $U_2=28<U_1=36$，故检验统计量 $U=28$；

(5) 由于 $U>13$，故接受 H_0，即用A饮料的女运动员的赛跑时间与用B饮料的女运动员的赛跑时间无差别.

20.19 重新考虑习题20.18，这次抽取更大样本，重新做该试验，现在，$n_1=30, n_2=30$，试验结果发现 $R_1=830, R_2=1,000$，且无等秩情形. 给定 $\alpha=0.05$，利用临界值决策规则，对零对假设 H_0：用A饮料的女运动员的赛跑时间与用B饮料的女运动员的赛跑时间相同，做双侧Mann-Whitney U 检验.

解 ☞　(1) H_0：用A饮料的女运动员的赛跑时间分布与用B饮料的女运动员赛跑时间分布相同，H_1：用A饮料的女运动员的赛跑时间分布与用B饮料的女运动员的赛跑时间分布不同；

(2) $\alpha=0.05$；

(3) 由于 $n_1>20, n_2>20$，故采用与 Z 变换有关的大样本的Mann-Whitney U 检验，由表A.5，对于双侧检验，临界值 $z_{0.05}=1.96$，决策规则为：若 $z^*<-1.96$，则拒绝 H_0；

(4) 将 $n_1=30, n_2=30, R_1=830$ 代入(20.6)式有

$$U_1=(30)(30)+\frac{(30)(30+1)}{2}-830=535$$

将 $n_1=30, n_2=30, R_2=1,000$ 代入(20.7)式，有

$$U_2=(30)(30)+\frac{(30)(30+1)}{2}-1,000=365$$

由于 $U_2=365<U_1=535$，故取 $U=365$，由于没有秩相等的情形发生，故不必采用修正因子，将 $U=365, n_1=30, n_2=30$ 代入(20.8)式，得

$$z^*=\frac{365-\frac{(30)(30)}{2}}{\sqrt{\frac{(30)(30)(30+30+1)}{12}}}=-1.2567$$

为确定 P 值，利用(16.4)式：$P=2[P(Z\geqslant(z^*=a)\mid H_0\text{ 为真})]$. 由表A.5知，立于区间 $0\leqslant z\leqslant1.26$ 上方区域的面积为0.3962，由于 Z 分布位于原点右侧的区域面积为0.5，故

$$P=2[P(Z\geqslant1.26\mid H_0\text{ 为真})]=2(0.5-0.3962)=0.2076$$

(5) 由于 $z^*>-1.96$，故接受 H_0，这一结论可由下述事实得以确认：在零假设为真时，得到 $z^*=-1.257$ 的概率大于0.05，因而A饮料效果与B饮料效果不同.

20.20 考虑习题20.19中的试验，但这次试验有秩相等的情形：2个8，3个10，5个11，分析这些等秩的出现对接受 H_0 这一决策的影响.

解 ☞　为分析等秩的影响，需要利用(20.10)式求出 z^*，并与(20.8)式得到的结果比较. 首先由(20.9)式得

$$\sum T_i=\sum\frac{t_i^3-t_i}{12}=\frac{2^3-2}{12}+\frac{3^3-3}{12}+\frac{5^3-5}{12}=12.5$$

将 $\sum T_i=12.5$ 代入(20.10)式

$$z^*=\frac{365-\frac{(30)(30)}{2}}{\sqrt{\left(\frac{(30)(30)}{(30+30)^2-(30+30)}\right)\left(\frac{(30+30)^3-(30+30)}{12}-12.5\right)}}=-1.2571$$

由于某些秩相等，使得 z^* 由 -1.2567 减少到 -1.2571，而 -1.2571 仍大于临界值 -1.96，而且得到

该值的概率为 $P=0.208$，从而结论与没有等秩的习题 20.19 相同.

由此可以知道，某些秩相等对于 z 有较小影响，但不足以改变我们接受 H_0 的决策，在零假设为真时，得到该样本值的概率至少在第三位小数上几乎保持不变.

多样本检验：k 个独立样本的 Kruskal-Wallis H 检验

20.21 一位环保人员想知道某种鹿居留于一种栖息地的时间是否多于居留于其他栖息地的时间，为此，他计算出松林地带的四块区域中每公顷鹿的个数，杉树地带的四块区域中每公顷鹿的个数，白杨树地带的三块区域中鹿的个数，调查结果如下：在松林地带，分别为 12,11,8,6；在杉树地带，分别为 8,10,4,5；在白杨树地带，分别为 13,7,9. 给定 $\alpha=0.05$，利用临界值决策规则，对零假设 H_0：三种森林地带鹿的密度相同，做 Kruskal-Wallis 检验.

解 (1) H_0：三种森林地带鹿的密度完全相同，H_1：三种森林地带鹿的密度不完全相同；

(2) $\alpha=0.05$；

(3) 由于 $k=3$，且 n 的值均小于 6，故采用小样本的 Kruskal-Wallis H 检验. 由表 A.14，对于 $n_1=4, n_2=4, n_3=3$，有 $H_{0.05}=5.598$，决策规则为：若 $H>5.598$，则拒绝 H_0；

(4) 对于每个处理计算 R，从而求得检验统计量 H，首先将三个处理的观测值合并，并按从小到大排序，记下其秩，若有相等的观测值，则取其秩平均值作为秩. 记第 i 个处理的秩和为 R_i，计算结果如表 20.23 所示.

表 20.23

松林地带		杉树地带		白杨地带	
鹿的个数	秩	鹿的个数	秩	鹿的个数	秩
12	10.0	8	5.5	13	11.0
11	9.0	10	8.0	7	4.0
8	5.5	4	1.0	9	7.0
6	3.0	5	2.0		
$n_1=4$	$R_1=27.5$	$n_2=4$	$R_2=16.5$	$n_3=3$	$R_3=22.0$

由于 11 个观测值中只有两个相等的观测值，不足 35%，故不必采用修正因子，利用(20.11)式得

$$H=\frac{12}{11(11+1)}\left(\frac{(27.5)^2}{4}+\frac{(16.5)^2}{4}+\frac{(22.0)^2}{3}\right)-3(11+1)=2.0417$$

查表 A.14 知，得到该 H 值的概率为 $P>0.10$；

(5) 由于 $H<5.598$，故接受 H_0，这一结论可由下述事实得以确认：在零假设为真时，得到 $H=2.0417$ 的概率大于 0.05，因此，三种森林地带鹿的密度无差别.

20.22 一位地质学家想知道某种岩石结构中的五块露岩中钾碱含量是否相同，为此，从每块露岩中抽取 8 个样本，测量其钾碱含量，结果如下：对于露岩 A，钾碱含量分别为 2.70, 7.50, 5.00, 2.25, 2.80, 6.00, 4.35, 3.00；对于露岩 B，钾碱含量分别为 4.10, 7.40, 6.65, 5.70, 5.30, 7.20, 6.30, 6.95；对于露岩 C，钾碱含量分别为 3.50, 3.20, 3.10, 4.20, 3.15, 4.30, 4.45, 5.25；对于露岩 D，钾碱含量分别为 6.85, 4.00, 4.25, 4.40, 6.55, 3.25, 2.20, 1.80；对于露岩 E，钾碱含量分别为 6.25, 6.05, 4.10, 4.05, 6.90, 6.45, 6.70, 5.30. 给定 $\alpha=0.01$，利用临界值决策规则，对零假设 H_0：五块露岩中钾碱含量均相同，做 Kruskal-Wallis 检验.

解 (1) H_0：五块露岩中钾碱含量均相同，H_1：五块露岩中钾碱含量不完全相同；

(2) $\alpha=0.01$；

(3) 由于 $k=3$，且 $n_i>5$，$(i=1,2,\cdots,5)$，故采用 χ^2 分布求临界值，由表 A.7，对于 $v=5-1=4$，临界值 $\chi^2_{0.01,4}=13.28$，决策规则为：若 $H>13.28$，则拒绝 H_0；

(4) 对于每个处理计算 R，从而求得检验统计量 H. 首先，将五个处理的所有观测值合并，按从

小到大排序,记下其所在位置,若有相等观测值的,则取其位置的平均作为秩,记第 i 个处理的秩和为 R_i,计算结果如表 20.24 所示.

表 20.24

钾碱含量(ppm)									
露岩 A		露岩 B		露岩 C		露岩 D		露岩 E	
含量	秩	含量	秩	含量	秩	含量	秩	含量	秩
2.70	4.0	4.10	14.5	3.50	11.0	6.85	35.0	6.25	29.0
7.50	40.0	7.40	39.0	3.20	9.0	4.00	12.0	6.05	28.0
5.00	22.0	6.65	33.0	3.10	7.0	4.25	17.0	4.10	14.5
2.25	3.0	5.70	26.0	4.20	16.0	4.40	20.0	4.05	13.0
2.80	5.0	5.30	24.5	3.15	8.0	6.55	32.0	6.90	36.0
6.00	27.0	7.20	38.0	4.30	18.0	3.25	10.0	6.45	31.0
4.35	19.0	6.30	30.0	4.45	21.0	2.20	2.0	6.70	34.0
3.00	6.0	6.95	37.0	5.25	23.0	1.80	1.0	5.30	24.5
$n_1=8$	$R_1=126$	$n_2=8$	$R_2=242$	$n_3=8$	$R_3=113$	$n_4=8$	$R_4=129$	$n_5=8$	$R_5=210$

由于 40 个观测值中只有 4 个观测值相同,仅占 10%,故不必采用修正因子,利用(20.11)式可得

$$H=\frac{12}{40(41)}\left(\frac{(126)^2}{8}+\frac{(242)^2}{8}+\frac{(113)^2}{8}+\frac{(129)^2}{8}+\frac{(210)^2}{8}\right)-3(40+1)=12.3201$$

查表 A.7 知,得到该 H 值的概率为 $0.01<P<0.025$;

(5) 由于 $H<13.28$,故在 0.01 的显著水平下接受 H_0,这一结论可由下述事实得以确认:在零假设为真时,得到 $H=12.3201$ 的概率大于 0.010,因此,认为五块露岩中的钾碱含量并无差别.

20.23 重新考虑习题 20.22 中的试验,假定这位地质学家对所得观测值采用四舍五入法进行处理,结果使得有两个观测值为 3.2,三个观测值为 4.1,两个观测值为 4.3,两个观测值为 4.4,三个观测值为 5.3,两个观测值为 6.3,两个观测值为 6.7,利用习题 20.22 得到 R_i 值,采用修正因子,给定 $\alpha=0.01$,对零假设 H_0:五块露岩中钾碱含量完全相同,做 Kruskal-Wallis 检验.

解 (1) H_0:五块露岩中钾碱含量完全相同,H_1:五块露岩钾碱含量不完全相同;

(2) $\alpha=0.01$;

(3) 由于 $k>3,n_i>5,(i=1,2,\cdots,5)$,故采用大样本的 Kruskal-Wallis H 检验,由表 A.7,对于自由度 $v=5-1=4,\alpha=0.01$,临界值 $\chi^2_{0.01,4}=13.28$,决策规则为:若 $H>13.28$,则拒绝 H_0;

(4) 对于每个处理求出 R,从而得到检验统计量 H 的值,由习题 20.22 知 $R_1=126,R_2=242,R_3=113,R_4=129,R_5=210,n_1=n_2=n_3=n_4=n_5=8$.由于 40 个观测值中相等的值占总数的 40%,故应采用修正因子来计算 H,由(20.12),(20.13)式得:

$$H=\frac{\frac{12}{40(41)}\left(\frac{(126)^2}{8}+\frac{(242)^2}{8}+\frac{(113)^2}{8}+\frac{(129)^2}{8}+\frac{(210)^2}{8}\right)-3(40+1)}{1-\frac{(2^3-2)+(3^3-3)+(2^3-2)+(2^3-2)+(3^3-3)+(2^3-2)+(2^3-2)}{(40)^3-40}}$$

$$=12.3352$$

查表 A.7 知,得到该 H 值的近似概率为 $0.01<P<0.025$;

(5) 修正因子将 H 值从 12.3201 增大到 12.3352,但并未达到临界值 13.28,因而,在零假设为真时,得到该 H 值的近似概率等于无修正因子时得出的概率,故我们接受零假设 H_0,即:五块露岩的钾碱含量完全相同.

Spearman 秩相关检验

20.24 某大学化学系决定拨出部分科研经费资助三名本科生,有 10 名本科生申请,该校安排

两名教授根据学生的研究计划,来评定学生的级别,10名学生假定为A,B,C,…,J;每位教授按从1(最好)到10(最差)给他们评定秩,A教授的评定情况为2,5,7,1,3,8,9,6,10,4;B教授的评定情况为1,3,8,2,5,7,10,4,9,6.给定$\alpha=0.05$,利用临界值决策规则,对假设H_0:两位教授的评价没有任何关系,和H_1:两位教授的评价有单调递增的关系,做Spearman秩相关检验.

解 (1) H_0:两位教授的评价没有任何关系,H_1:两位教授的评价有单调递增的关系;

(2) $\alpha=0.05$;

(3) 由于评价总数为10对,小于30,故采用表A.15求临界值.对于给定$\alpha=0.05$,$n=10$的右侧检验,临界值为0.564,决策规则为:若$r_S>0.564$,则拒绝H_0;

(4) 由于没有秩相等的情形,故采用(20.14)式求相关系数,为此,对每对观测值求秩的差,然后平方得到d_i^2,计算结果见表20.25.

表20.25

学生	A教授的评价	B教授的评价	d_i	d_i^2
A	2	1	+1	1
B	5	3	+2	4
C	7	8	−1	1
D	1	2	−1	1
E	3	5	−2	4
F	8	7	+1	1
G	9	10	−1	1
H	6	4	+2	4
I	10	9	+1	1
J	4	6	−2	4
				$\sum=22$

将这些平方值求和得$\sum=22$,注意到$n=10$,由(20.14)式得

$$r_S=1-\frac{6(22)}{10(10^2-1)}=0.8667$$

查表A.15知,得到该相关系数的近似概率为$0.001<P<0.0025$;

(5) 由于$r_S>0.564$,故拒绝H_0,接受H_1,这一结论可由下述事实得以确认:得到$r_S=0.8667$的概率小于0.05,因此,两个教授的评价相似,有单调递增的关系.

20.25 由于习题20.24中的研究经费的成功运作,使得第二次有32人申请.再次由这两位教授为这些申请者评定秩,假定没有等秩的情形,整理数据后,得出所有秩差的平方和为3,010,给定$\alpha=0.01$,利用临界值决策规则,对假设H_0:两位教授的评价没有任何关系,和H_1:两位教授的评价有单调递增关系,做Spearman秩相关检验.

解 (1) H_0:两个评价总体没有任何关系,H_1:两个评价总体有单调递增的关系;

(2) $\alpha=0.01$;

(3) 由于共有32对秩,大于30,故作Z变换,利用Z分布求临界值,由表A.5,对于右侧检验,临界值$z_{0.01}=2.33$,决策规则为:若$z^*>2.33$,则拒绝H_0;

(4) 由于没有等秩情形,故采用(20.14)式求相关系数,已知$\sum_{i=1}^{32}d_i^2=3010$,$n=32$,于是有

$$r_S=1-\frac{6(3,010)}{32(32^2-1)}=0.4483$$

利用(20.15)式将r_S转换成z

$$z^* = 0.4483\sqrt{32-1} = 2.4960 \approx 2.50$$

查表 A.5 知,得到该样本值的概率为 $P = 0.5 - 0.4938 = 0.0062$;

(5) 由于 $z^* > 2.33$,故在 0.01 的显著水平下拒绝 H_0,这一结论可由下述事实得以确认:得到该样本相关系数的概率小于 0.01,因而两组评价有单调递增的关系.

20.26 某大学的体育主任想知道该校的篮球队在联赛结束时的名次是否与首发阵容中五名参赛人员的平均身高有关,该校去年联赛共有 8 支球队,名次分别为 1 至 8,相应的首发阵容中五名运动员平均身高分别为 6 英尺 5.5 英寸,6 英尺 5.0 英寸,6 英尺 6.5 英寸,6 英尺 3.8 英寸,6 英尺 7.2 英寸,6 英尺 4.4 英寸,6 英尺 6.2 英寸,6 英尺 6.1 英寸,给定显著水平 $\alpha = 0.05$,利用临界值决策规则,对假设 H_0:球队名次与队员平均身高的名次无关,和 H_1:球队名次与队员平均身高名次有单调增的关系,做 Spearman 秩相关检验.

解 (1) H_0:球队名次与队员平均身高无关,H_1:球队名次与队员平均身高的名次有单调增的关系;

(2) $\alpha = 0.05$;

(3) 由于样本对小于 30,故利用表 A.15 求临界值,对于 $\alpha = 0.05$, $n = 8$ 的右侧检验,决策规则为:若 $r_S > 0.643$,则拒绝 H_0;

(4) 由于没有出现等秩情形,故采用(20.14)式计算相关系数,首先,计算 r_S,由于球队名次已确定,故只需将队员平均身高按从大到小排序,然后对每个球队求出 d_i 和 d_i^2,计算结果见表 20.26.

表 20.26

球队名次	平均身高	身高的秩	d_i	d_i^2
1	6 英尺 5.5 英寸	5	−4	16
2	6 英尺 5.0 英寸	6	4	16
3	6 英尺 6.5 英寸	2	+1	1
4	6 英尺 3.8 英寸	8	4	16
5	6 英尺 7.2 英寸	1	+4	16
6	6 英尺 4.4 英寸	7	−1	1
7	6 英尺 6.2 英寸	3	+4	16
8	6 英尺 6.1 英寸	4	+4	16
				$\sum = 98$

由于 $\sum = 98$, $n = 8$,利用(20.14)式,有

$$r_S = 1 - \frac{6(98)}{8(8^2-1)} = -0.1667$$

查表 A.15 知,得到该样相关系数的概率为 $P > 0.25$;

(5) 由于 $r_S < 0.643$,故在 0.05 的显著水平下接受 H_0,这一结论可由下述事实得以确认:在零假设为真时,得到 $r_S = -0.1667$ 的概率大于 0.05,因此,赛季末球队的名次与首发阵容中五名队员的平均身高无关.

补充习题

χ^2 拟合优度检验

20.27 一枚均匀骰子连续掷 60 次,由于每个面朝上的概率均为 1/6,故某个指定面的期望频数为 $\left(\frac{1}{6}\right)(60) = 10$,观测频率(按从 1 点至 6 点顺序排列)分别为 7,7,11,12,13,10. 给定显著性水平 $\alpha = 0.05$,对零假设 H_0:观测频数服从均匀分布,做 χ^2 检验. 给出检验统计量的样本值,说明接受 H_0 还是拒绝 H_0,并给出在 H_0 为真时,得到该 χ^{2*} 值的近似概率.

答案：$\chi^{2*}=3.20$，由于 $\chi^{2*}<\chi^2_{0.05,5}=11.07$，故接受 H_0．$0.500<P<0.900$

20.28 某商店出售五种类型的牙刷，商店经理想知道这五种牙刷是否以相同频数售出，为此他随机抽取销售量为 500 的一个样本，并记录品牌，如果五种牙刷是以相同频数售出，那么每个品牌的牙刷的售出期望频数应为 100，已知观测频数分别为 81，135，97，90，97．给定显著性水平 $\alpha=0.05$，对零假设 H_0：观测频数服从均匀分布，做 χ^2 检验．给出检验统计量的样本值，说明接受 H_0，还是拒绝 H_0，并给出在 H_0 为真时得到该 χ^{2*} 值的近似概率．

答案：$\chi^{2*}=17.04$，由于 $\chi^{2*}>\chi^2_{0.05,4}=9.49$，故拒绝 H_0．$P<0.005$

20.29 一位基因学家想知道某种特征遗传的基因型是否为简单的孟德尔隐性遗传方式，如果这一结论成立那么两个杂合体(每个均有这种基因的两种形式)交配后所产生的后代中的这种基因型出现的比例将为 1∶2∶1．他做此试验，结果如下：200 个后代中该基因型的比例为 35∶95∶70．给定显著性水平 $\alpha=0.01$，对零假设 H_0：观测频数比例为 1∶2∶1，作 χ^2 检验．给出检验统计量的样本值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该 χ^{2*} 值的近似概率．

答案：$\chi^{2*}=12.75$，由于 $\chi^{2*}>\chi^2_{0.01,2}=9.21$，故拒绝 H_0．$P<0.005$

独立性的 χ^2 检验：列联表分析

20.30 一位大学的招生人员想知道学生报考省内学校或省外学校是否与其家庭状况(单亲家庭或双亲家庭)有关，他从某大城市随机抽取 60 名学生，调查结果表明，其中有 25 人来自单亲家庭，35 人来自双亲家庭；在单亲家庭的学生中，15 人报考省内学校，10 人报考省外学校，而在双亲家庭的学生中，27 人报考省内学校，8 人报考省外学校．给定显著性水平 $\alpha=0.05$，采用修正因子，对零假设 H_0：学生报考省内或省外学校与学生来自单亲或双亲家庭无关，做 χ^2 检验．给出检验统计量的样本值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时，得到该 χ^{2*} 值的近似概率．

答案：$\chi^{2*}=1.306$，由于 $\chi^{2*}<\chi^2_{0.05,1}=3.84$，故接受 H_0．$0.100<P<0.500$

20.31 一位中学教务处长想知道城市父母和乡村父母对于延长学生在校时间是否持不同态度，他随机抽取两组样本，200 位父母来自城市，300 位父母来自乡村，询问他们对于此事所持态度：赞成、反对或没有看法，调查结果为：城市父母中，123 人支持，36 人反对，41 人没有看法；乡村父母中，145 人支持，85 人反对，70 人没有看法．给定显著性水平 $\alpha=0.01$，对零假设 H_0：对于延长学生在校时间所持态度与父母住在城市或农村无关，做 χ^2 检验．给出检验统计量的样本值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该 χ^{2*} 值的近似概率．

答案：$\chi^{2*}=9.610$，由于 $\chi^{2*}>\chi^2_{0.01,2}=9.21$，故拒绝 H_0．$0.005<P<0.010$

20.32 某县欲在两个主要城市之间建一条高速公路，一位新闻记者想知道居民对此计划所持态度是否依赖于他们的政治倾向，他随机选取 1,000 人做调查，结果如下：(1)被调查的 420 位民主党人士中，有 211 人赞成，144 人反对，65 人持中立态度；(2)被调查的 440 位共和党人士中，有 248 人赞成，141 人反对，51 人持中立态度；(3)被调查的 140 位无党派人士中，76 人赞成，40 人反对，24 人持中立态度．给定显著性水平 $\alpha=0.05$，对零假设 H_0：对修建高速公路的态度与其政治立场无关，做 χ^2 检验．给出检验统计量的样本值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该 χ^{2*} 值的近似概率．

答案：$\chi^{2*}=6.096$，由于 $\chi^{2*}<\chi^2_{0.05,4}=9.49$，故接受 H_0．$0.100<P<0.500$

k 个二项比率齐性的 χ^2 检验

20.33 一位文具店老板欲购入一批铅笔，共四种品牌：A，B，C 和 D．他想知道对于这四种品牌，不受欢迎的铅笔所占比率是否相同．为此，她随机抽取 1,000 名顾客，每种铅笔 250 人，询问他们认为该品牌的铅笔受欢迎还是不受欢迎，调查结果如下：70 人认为 A 铅笔不受欢迎，40 人认为 B 铅笔不受欢迎，30 人认为 C 铅笔不受欢迎，60 人认为 D 铅笔不受欢迎．给定显著性水平 $\alpha=0.01$，对零假设 H_0：这四种品牌铅笔不受欢迎的比率相同，做 χ^2 检验．给出检验统计量的样本值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该 χ^{2*} 值的近似概率．

答案：$\chi^{2*}=25.00$，由于 $\chi^{2*}>\chi^2_{0.01,3}=11.34$，故拒绝 H_0．$P<0.005$

20.34 一位社会学家想知道参加男童子军的青少年的比率是否随年龄而变化，他从年龄在 13 岁至 17 岁的五个年龄段的青少年中随机抽取样本，每个年龄段 100 个人，并记录下其中参加童子军的人数，结果如下：13 岁的有 26 人，14 岁的有 29 人，15 岁的有 32 人，16 岁的有 31 人，17 岁的有 30 人．给定显著性水平 $\alpha=0.05$，对零假设 H_0：五个年龄段的男孩中，参加童子军的人数比率相同，做 χ^2 检验．给出检验统计量的样本值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该 χ^{2*} 值的概率．

答案:$\chi^{2*}=1.017$,由于 $\chi^{2*}<\chi^2_{0.05,4}=9.49$,故接受 H_0.$0.900<P<0.950$

20.35 一位中学行政人员想帮助该校学生通过驾车考试,她对四所驾校随机抽样,每所驾校30位学员,调查其初次参加驾车考试就通过的人数,结果如下:A校有18人;B校有26人;C校有17人;D校有19人.给定显著性水平 $\alpha=0.05$.对零假设 H_0:四所驾校初次参加驾车考试即能通过的比例相同,做 χ^2 检验.给出检验统计量的样本值,说明接受 H_0 还是拒绝 H_0,并给出在 H_0 为真时得到该 χ^{2*} 值的近似概率.

答案:$\chi^{2*}=7.500$,由于 $\chi^{2*}<\chi^2_{0.05,3}=7.81$,故拒绝 H_0.$0.050<P<0.100$

单样本检验:Wilcoxon 符号秩检验

20.36 已知某蜥蜴总体从口鼻部到排泄口长度的中位数为54毫米(假定总体数量相当大).该群体中的一部分蜥蜴迁徙到某一小岛上后经过几代的繁衍生殖,形成一个较小的蜥蜴总体,我们想知道这两个总体的蜥蜴身体长度的中位数是否相同.为此,从较小的蜥蜴总体中随机抽取容量为25的一个样本,结果发现:$W_+=98$,$W_-=-227$,给定 $\alpha=0.05$,对假设 $H_0:\tilde{\mu}=\tilde{\mu}_0$ 和 $H_1:\tilde{\mu}\neq\tilde{\mu}_0$,做双侧 Wilcoxon 检验.给出检验统计量的样本值,说明接受 H_0 还是拒绝 H_0,并给出在 H_0 为真时得到该 W 值的近似概率.

答案:$W=98$,由于 $W>90$,故接受 H_0.$0.05<P<0.10$

20.37 一位鞋业连锁店的经理计划在某镇开设一家分店,当地政府告诉他该镇每个家庭收入的中位数为\$3,600,这位经理认为这个估计过高,于是随机调查8户家庭的收入,调查结果如下:\$20,000,\$37,000,\$36,500,\$36,400,\$21,000,\$32,600,\$30,000,\$37,000.给定显著性水平 $\alpha=0.05$,对假设 $H_0:\tilde{\mu}=\tilde{\mu}_0$,和 $H_1:\tilde{\mu}<\tilde{\mu}_0$,做单侧 Wilcoxon 检验.给出检验统计量的样本值,说明接受 H_0 还是拒绝 H_0,并给出在 H_0 为真时得到该样本值 W 的近似概率.

答案:$W=10$,由于 $W>6$,故接受 H_0.$P>0.05$

20.38 高中生智商的中位数为100.一位教师想知道她所教班级学生的智商中位数是否也为100.为此她从其班中随机抽取12名学生,测他们的智商分数,结果如下:110,98,120,130,110,97,95,125,135,128,94,96.给定 $\alpha=0.05$,对假设 $H_0:\tilde{\mu}=\tilde{\mu}_0$ 和 $H_1:\tilde{\mu}\neq\tilde{\mu}_0$,做双侧 Wilcoxon 检验.给出检验统计量的样本值,说明接受 H_0 还是拒绝 H_0,并给出在 H_0 为真时得到该样本值 W 的近似概率.

答案:$W=15$,由于 $W>14$,故接受 H_0.$0.05<P<0.10$

两样本检验:独立样本的 Wilcoxon 符号秩检验

20.39 一文秘培训班开设了一门打字课程,每年都做调查看看该课程是否真正提高了学员们每分钟的打字数.培训班主管随机抽取15名学员,记录下其打字速度(个/分钟),结果如下:(100,120);(105,100);(90,96);(120,118);(110,128);(95,94);(98,121);(97,135);(86,93);(95,87);(99,111);(115,140);(115,112);(120,130);(120,109);其中每对数字的第一个表示培训前的打字速度,第二个表示培训后的打字速度.给定显著性水平 $\alpha=0.05$.对于假设 $H_0:\tilde{\mu}_1=\tilde{\mu}_2$,和 $H_1:\tilde{\mu}_1<\tilde{\mu}_2$,做小样本的 Wilcoxon 检验.给出检验统计量的样本值,说明接受 H_0 还是拒绝 H_0,并给出在 H_0 为真时得到该样本值 W 的近似概率.

答案:$W=26$,由于 $W<30$,故拒绝 H_0.$0.025<P<0.050$

20.40 续习题20.39,培训班主管想在广告中对打字课做一强有力的论证,为此,她随机抽取50人的更大样本,记录下他们培训前和培训后的打字速度(单位:个/分钟),分别记为样本1和样本2,求得速度的差,将这些差的得分排序,记录下其秩,并将差的得分的"+"或"-"号加到秩前,最后求得 $W_+=365$,$W_-=-910$.给定 $\alpha=0.01$,对于假设 $H_0:\tilde{\mu}_1=\tilde{\mu}_2$,和 $H_1:\tilde{\mu}_1<\tilde{\mu}_2$,做大样本的 Wilcoxon 检验.给出检验统计量的样本值,说明接受 H_0,还是拒绝 H_0,并给出在 H_0 为真时,得到该 z^* 值的近似概率.

答案:$z^*=-2.631$,由于 $z^*<-z_{0.01}=-2.33$,故拒绝 H_0.$P\approx0.004$

20.41 一位汽车维修部门经理近日购入两种车胎:A和B,想知道这两种车胎的质量是否相同(根据损坏前走过的英里数判断),他随机选取8部车,从车胎A,B中随机地选一个装在左车轮上,另一个装在右车轮上,然后记录下每个车胎在损坏前走过的英里数,结果如下(A车胎的值列在前面):(41,000;40,000);(38,000;40,000);(38,500;40,000);(43,000;40,000);(42,500;41,400);(41,500;42,300);(40,100;40,200);(39,600;41,300).给定 $\alpha=0.05$,对假设 $H_0:\tilde{\mu}_1=\tilde{\mu}_2$ 和 $H_1:\tilde{\mu}_1\neq\tilde{\mu}_2$,做双侧 Wilcoxon 检验.给出检验统计量的样本值,说明接受 H_0 还是拒绝 H_0,并给出在 H_0 为真时得到该 W 值的近似概率.

答案：$W=15$，由于 $W>4$，故接受 H_0. $P>0.10$

两样本检验：独立样本的 Mann-Whitney U 检验

20.42　一位教育学家研究了两种教四年级学生数学的新方法，为判断方法 A 是否好于方法 B，他随机选取 6 名学生用 A 方法教课（样本 1），8 名学生用方法 B 教课（样本 2），16 星期后，他对两组学生安排一次完全相同的考试，并记录下他们正确地解答所有题的时间，结果如下（按分钟记）：对于样本 A，时间分别为：50，45，30，42，36，33；对于样本 B，时间分别为 52，51，47，35，40，38，53，34. 给定 $\alpha=0.05$，对于假设 H_0：总体 A 的得分的分布与总体 B 的得分的分布相同，和 H_1：总体 A 的得分的分布小于总体 B 的得分的分布，做单侧 Mann-Whitney 检验. 给出检验统计量的样本值，并说明接受 H_0 还是拒绝 H_0.

答案：$U=15$，由于 $U>10$，故接受 H_0

20.43　习题 20.42 中的教育家随机选取更大样本（每种方法 40 人）重新做试验，结果发现 $R_1=1,250$，$R_2=1,990$，且没有花费时间相同的情形. 给定 $\alpha=0.01$，对于假设 H_0：总体 A 的得分的分布与总体 B 的得分的分布相同，和 H_1：总体 A 的得分的分布小于总体 B 的得分的分布，做单侧 Mann-Whitney U 检验. 给出检验统计量的样本值，并说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该 z^* 值的概率.

答案：$z^*=-3.560$，由于 $z^*<-2.33$，故拒绝 H_0. $P=0.0002$

20.44　重新考虑习题 20.43 中的试验，但这时有相同秩出现：三个秩为 2，五个秩为 5，八个秩为 10. 给出检验统计量的样本值，并说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该 z^* 值的概率.

答案：$z^*=-3.563$，由于 $z^*<-2.33$，故拒绝 H_0. $P=0.0002$

多样本检验：k 个独立样本的 Kruskal-Wallis H 检验

20.45　一位八年级的数学教师采用三种不同方法教几何，对于每种教学方法，随机选择 4 名学生（共 12 名学生），两学期后参加综合考试，根据他们的成绩来评价三种教学方法，试验结果如下：对于方法 A，学生成绩分别为 60，55，78，66；对于方法 B，学生成绩分别为 61，74，60，63；对于方法 C，学生分别为 72，77，82，80. 给定 $\alpha=0.05$，对零假设 H_0：对于三种教学法，学生成绩分布相同，做 Kruskal-Wallis H 检验. 给出检验统计量的样本值，并说明接受 H_0 还是拒绝 H_0.

答案：$H=4.875$，由于 $H<H_{0.05}=5.692$，故接受 H_0

20.46　习题 20.45 中的数学教师重新试验，这次对于每种教学法随机抽取 10 名学生，试验结果为：对于方法 A，成绩分别为：50，60，64，67，63，75，81，55，51，53；对于方法 B，成绩分别为：58，43，49，65，74，68，71，70，57，66；对于方法 C，成绩分别为 70，78，80，76，86，74，72，73，83，84. 给定 $\alpha=0.01$，对零假设 H_0：三种教学方法的学生成绩分布相同，做 Kruskal-Wallis H 检验. 给出检验统计量的样本值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该样本值的概率.

答案：$H=13.659$，由于 $H>\chi^2_{0.01,2}=9.21$，故拒绝 H_0. $P<0.005$

20.47　一位高尔夫球教练想知道高尔夫球运动员的年龄与他们在巡回比赛中赢得的奖金是否有关系，为此，他在三个年龄段中的每个年龄段随机选取 4 人，得到结果如下：在 25 岁以下，奖金为 \$7,535，\$10,000，\$13,000，\$12,000；26 岁至 30 岁之间，奖金分别为 \$25,000，\$18,500，\$13,500，\$20,000；30 岁以上，奖金分别为 \$21,900，\$30,000，\$33,000，\$28,000. 给定 $\alpha=0.05$，对零假设 H_0：三个年龄段的人赢得奖金的分布相同，做 Kruskal-Wallis H 检验. 给出检验统计量的样本值. 并说明接受 H_0 还是拒绝 H_0.

答案：$H=9.269$，由于 $H>H_{0.05}=5.692$，故拒绝 H_0

20.48　（续习题 20.47）该教练抽取更大的样本，重做他的研究，这次他在每个年龄段随机抽取 7 人，调查结果如下：25 岁以下，奖金分别为：\$28,000，\$12,000，\$15,000，\$15,500，\$12,500，\$14,000，\$14,500；26 岁至 30 岁之间，奖金分别为：\$30,000，\$34,000，\$16,000，\$18,000，\$21,000，\$16,500，\$17,000；30 岁以上，奖金分别为：\$20,000，\$25,000，\$28,000，\$35,000，\$29,500，\$42,500，\$41,000. 给定 $\alpha=0.01$，对零假设 H_0：三个年龄段的运动员赢得奖金的分布相同，做 Kruskal-Wallis H 检验. 给出检验统计量的样本值，说明接受 H_0 还是拒绝 H_0，并给出在 H_0 为真时得到该 χ^{2*} 值的近似概率.

答案：$H=11.668$，由于 $H>\chi^2_{0.01,2}=9.21$，故拒绝 H_0. $P<0.005$

Spearman 秩相关检验

20.49　某专家组按两个标准对 10 个城镇评定秩：生活水平（最好水平的秩为 1）和犯罪率（犯罪率最高的秩

为1),每一个变量均没有等秩情形.将每一个城镇的犯罪率的名次减去其生活水平的名次,得:−7,−7,−3,1,0,−2,0,4,8,8.给定 $\alpha=0.05$,对零假设 H_0:生活水平的名次与犯罪率的名次无关,做双侧 Spearman 秩相关检验.给出检验统计量的样本值,并说明接受 H_0 还是拒绝 H_0.

答案:$r_S=-0.552$,由于 $r_S>-0.648$,故接受 H_0

20.50　一位攀岩教练认为:由于业余攀岩队员较多,从而导致难度较低的攀岩比难度较高的攀岩发生更多的事故.他记录了过去5年内12次攀岩运动发生的事故数,按攀岩难度从1(最小难度)到12(最大难度)排秩,并按攀岩难度给出事故次数,结果分别为:2(难度最小),3,5,3,4,3,3,4,4,2,1,1(难度最大).给定 $\alpha=0.05$,对假设 H_0:事故发生次数与攀岩难度无关,和 H_1:事故发生次数与攀岩难度之间有单调递减的关系,做单侧 Spearman 秩相关检验.给出检验统计量的样本值,并说明接受 H_0,还是拒绝 H_0.

答案:$r_S=-0.4371$,由于 $r_S>-0.503$,故接受 H_0

20.51　一旅行社想知道旅游者假期的消费与旅行的距离是否相关,为此随机调查了40位度假者,并将他们的旅行距离按从小到大排序,记录下其秩(从1到40,且无相等情形).再记算出他们的周平均消费,按从小到大排序,记录下秩(无相等情形),将每个度假者周平均消费的秩减去其旅行距离的秩得:−10,−29,−6,8,−8,+5,−10,+6,−9,+6,−4,−4,+6,−20,−4,+6,−8,+10,−18,−6,−6,−18,5,−5,+1,+3,−6,−23,+26,8,+9,−3,+27,+5,+21,+6,+17,+17,+7,+4.给定 $\alpha=0.01$,对假设 H_0:假期消费与旅行距离无关,和 H_1:假期消费与旅行距离之间有单调增加的关系,做单侧 Spearman 秩相关检验.给出检验统计量的样本值,说明接受 H_0 还是拒绝 H_0,并给出在 H_0 为真时得到该 z 值的近似概率.

答案:$r_S=0.4263$,$z^*=2.66$,由于 $z^*>2.33$,故拒绝 H_0.$P<0.004$

附　　录

表 A.3　累积二项概率

本表给出 $n=2,3,\cdots,10$, $p=0.01,0.05,\cdots,0.50$ 的累积概率

$$F(a)=\sum_{x=0}^{a}\left[f(x)=\binom{n}{x}p^{x}q^{n-x}\right]$$

为了求 $p>0.5$ 时的 $F(a)$,在适当的 n 处,找到 $n-(a+1)$ 行与 $1-p$ 列的交点,然后从 1 减去这个值即是.

n	a	p = 0.01	0.05	0.10	0.15	0.20	0.25	0.30	1/3	0.35	0.40	0.45	0.50
2	0	0.9801	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4444	0.4225	0.3600	0.3025	0.2500
	1	0.9999	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8889	0.8775	0.8400	0.7975	0.7500
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.9703	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2963	0.2746	0.2160	0.1664	0.1250
	1	0.9997	0.9927	0.9720	0.9393	0.8960	0.8437	0.7840	0.7407	0.7183	0.6480	0.5748	0.5000
	2	1.0000	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9630	0.9571	0.9360	0.9089	0.8750
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	0.9606	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1975	0.1785	0.1296	0.0915	0.0625
	1	0.9994	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5926	0.5630	0.4752	0.3910	0.3125
	2	1.0000	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8889	0.8735	0.8208	0.7585	0.6875
	3	1.0000	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9877	0.9850	0.9744	0.9590	0.9375
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	0.9510	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1317	0.1160	0.0778	0.0503	0.0312
	1	0.9990	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4609	0.4284	0.3370	0.2562	0.1875
	2	1.0000	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7901	0.7648	0.6826	0.5931	0.5000
	3	1.0000	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9547	0.9460	0.9130	0.8688	0.8125
	4	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9959	0.9947	0.9898	0.9815	0.9687
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	0.9415	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0878	0.0754	0.0467	0.0277	0.0156
	1	0.9985	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3512	0.3191	0.2333	0.1636	0.1094
	2	1.0000	0.9978	0.9841	0.9527	0.9011	0.8306	0.7443	0.6804	0.6471	0.5443	0.4415	0.3437
	3	1.0000	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8999	0.8826	0.8208	0.7447	0.6562
	4	1.0000	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9822	0.9777	0.9590	0.9308	0.8906
	5	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9986	0.9982	0.9959	0.9917	0.9844
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

续表

n	a	p 0.01	0.05	0.10	0.15	0.20	0.25	0.30	1/3	0.35	0.40	0.45	0.50
7	0	0.9321	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0585	0.0490	0.0280	0.0152	0.0078
	1	0.9980	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2634	0.2338	0.1586	0.1024	0.0625
	2	1.0000	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.5706	0.5323	0.4199	0.3164	0.2266
	3	1.0000	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.8267	0.8002	0.7102	0.6083	0.5000
	4	1.0000	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9547	0.9444	0.9037	0.8471	0.7734
	5	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9962	0.9931	0.9910	0.9812	0.9643	0.9375
	6	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9995	0.9994	0.9984	0.9963	0.9922
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8	0	0.9227	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0390	0.0319	0.0168	0.0084	0.0039
	1	0.9973	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1951	0.1691	0.1064	0.0632	0.0352
	2	0.9999	0.9942	0.9619	0.8948	0.7969	0.6785	0.5518	0.4682	0.4278	0.3154	0.2201	0.1445
	3	1.0000	0.9996	0.9950	0.9786	0.9437	0.8862	0.8059	0.7414	0.7064	0.5941	0.4770	0.3633
	4	1.0000	1.0000	0.9996	0.9971	0.9896	0.9727	0.9420	0.9121	0.8939	0.8263	0.7396	0.6367
	5	1.0000	1.0000	1.0000	0.9998	0.9988	0.9958	0.9887	0.9803	0.9747	0.9502	0.9115	0.8555
	6	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9974	0.9964	0.9915	0.9819	0.9648
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9998	0.9993	0.9983	0.9961
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	0	0.9135	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0260	0.0207	0.0101	0.0046	0.0020
	1	0.9966	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.1431	0.1211	0.0705	0.0385	0.0195
	2	0.9999	0.9916	0.9470	0.8591	0.7382	0.6007	0.4628	0.3772	0.3373	0.2318	0.1495	0.0898
	3	1.0000	0.9994	0.9917	0.9661	0.9144	0.8343	0.7297	0.6503	0.6089	0.4826	0.3614	0.2539
	4	1.0000	1.000	0.9991	0.9944	0.9804	0.9511	0.9012	0.8552	0.8283	0.7334	0.6214	0.5000
	5	1.0000	1.0000	0.9999	0.9994	0.9969	0.9900	0.9747	0.9576	0.9464	0.9006	0.8342	0.7461
	6	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987	0.9957	0.9917	0.9888	0.9750	0.9502	0.9102
	7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9990	0.9986	0.9962	0.9909	0.9805
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9999	0.9997	0.9992	0.9980
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	0	0.9044	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0173	0.0135	0.0060	0.0025	0.0010
	1	0.9957	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.1040	0.0860	0.0464	0.0233	0.0107
	2	0.9999	0.9855	0.9298	0.8202	0.6778	0.5256	0.3828	0.2991	0.2616	0.1673	0.0996	0.0547
	3	1.0000	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5593	0.5138	0.3823	0.2660	0.1719
	4	1.0000	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7869	0.7515	0.6331	0.5044	0.3770
	5	1.0000	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9234	0.9051	0.8338	0.7384	0.6230
	6	1.0000	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9803	0.9740	0.9452	0.8980	0.8281
	7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9966	0.9952	0.9877	0.9726	0.9453
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9995	0.9983	0.9955	0.9893
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

来源：Wilfnd J. Dixon and Frank J. Massey, Introduction to Statistical Analysis(4th ed.), McGraw Hill, New York, 1983.

表 A.4　累积 Poisson 概率

本表给出 $\mu=0.001,\cdots,1.00$ 及 $\mu=1.1,\cdots,8.0$ 的累积概率

$$F(a)=\sum_{x\leq a}\left[f(x)=\frac{\mu^x e^{-\mu}}{x!}\right]$$

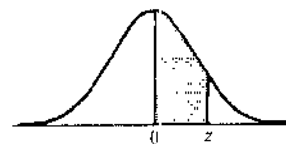
a	μ														
	0.001	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0	0.999	0.990	0.951	0.905	0.861	0.819	0.779	0.741	0.670	0.607	0.549	0.497	0.449	0.407	0.368
1	1.000	1.000	0.991	0.995	0.990	0.982	0.974	0.963	0.938	0.910	0.878	0.844	0.809	0.772	0.736
2			1.000	1.000	0.999	0.999	0.998	0.996	0.992	0.986	0.977	0.966	0.953	0.937	0.920
3					1.000	1.000	1.000	1.000	0.999	0.998	0.997	0.994	0.991	0.987	0.981
4									1.000	1.000	1.000	0.999	0.999	0.998	0.996
5												1.000	1.000	1.000	0.999
6															1.000

a	μ														
	1.1	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5	8.0
0	0.333	0.223	0.135	0.082	0.050	0.030	0.018	0.011	0.007	0.004	0.002	0.002	0.001	0.001	0.000
1	0.699	0.558	0.406	0.287	0.199	0.136	0.092	0.061	0.040	0.027	0.017	0.011	0.007	0.005	0.003
2	0.900	0.809	0.677	0.544	0.423	0.321	0.238	0.174	0.125	0.088	0.062	0.042	0.030	0.020	0.014
3	0.974	0.934	0.857	0.758	0.647	0.537	0.433	0.342	0.265	0.202	0.151	0.112	0.082	0.059	0.042
4	0.995	0.981	0.947	0.891	0.815	0.725	0.629	0.532	0.440	0.358	0.285	0.224	0.173	0.132	0.100
5	0.999	0.996	0.983	0.958	0.916	0.858	0.785	0.703	0.616	0.529	0.446	0.369	0.301	0.241	0.191
6	1.000	0.999	0.995	0.986	0.966	0.935	0.889	0.831	0.762	0.686	0.606	0.527	0.450	0.378	0.313
7		1.000	0.999	0.996	0.988	0.973	0.949	0.913	0.867	0.809	0.744	0.673	0.599	0.525	0.453
8			1.000	0.999	0.996	0.990	0.979	0.960	0.932	0.894	0.847	0.792	0.729	0.662	0.593
9				1.000	0.999	0.997	0.992	0.983	0.968	0.946	0.916	0.877	0.830	0.776	0.717
10					1.000	0.999	0.997	0.993	0.986	0.975	0.957	0.933	0.901	0.862	0.816
11						1.000	0.999	0.998	0.995	0.989	0.980	0.966	0.947	0.921	0.888
12							1.000	0.999	0.998	0.996	0.991	0.984	0.973	0.957	0.936
13								1.000	0.999	0.998	0.996	0.993	0.987	0.978	0.966
14									1.000	0.999	0.999	0.997	0.994	0.990	0.983
15										1.000	0.999	0.999	0.998	0.995	0.992
16											1.000	1.000	0.999	0.998	0.996
17													1.000	0.999	0.998
18														1.000	0.999
19															1.000

来源：Wilfrid J. Dixon and Frank J. Massey, Introduction to Statistical Analysis(4th ed.), McGraw-Hill, New York, 1983. 由 McGraw-Hill 公司授权翻印.

表 A.5　标准正态分布的面积

本表给出标准正态分布曲线以下,0 到 z 区间以上部分的面积(图中阴影区域),其中 z 表示标准正态变量 Z 的特定的正值.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2258	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2996	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

来源:Murray R. Spiegel, Schaum's Outline of Theory and Problems of Statistics(2nd ed.), McGraw-Hill, New York, 1988. 由 McGraw-Hill 公司授权翻印.

表 A.6　t 分布的临界值

本表给出 t 分布的正临界值 $t_{\alpha/2,\nu}$ 及 $t_{\alpha,\nu}$ ($\nu=1,\cdots,30,40,60,120,\infty$),表的使用说明见 14.20 节.

	$t_{\alpha/2}$					
	$t_{0.20/2}$	$t_{0.10/2}$	$t_{0.05/2}$	$t_{0.02/2}$	$t_{0.01/2}$	$t_{0.001/2}$
	t_α					
ν	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.0005}$
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	0.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

来源:Table III of Ronald A. Fisher and Frank Yates, Statistical Tables for Biological, Agricultural and Medical Research(6th ed.), Longman(朗文)Group Ltd., London, 1974.(早期由 Edinburgh 的 Oliver & Boyd 有限公司出版). 由 Pearson Education Limited 授权翻印.

表 A.7 χ^2 分布的临界值

本表给出 χ^2 分布的临界值 $\chi^2_{\alpha,\nu}$ ($\nu=1,\cdots,30,40,50,60,70,80,90,100$)，表的使用说明见 15.4 节.

ν	χ^2_α										
	$\chi^2_{0.995}$	$\chi^2_{0.990}$	$\chi^2_{0.975}$	$\chi^2_{0.950}$	$\chi^2_{0.900}$	$\chi^2_{0.500}$	$\chi^2_{0.100}$	$\chi^2_{0.050}$	$\chi^2_{0.025}$	$\chi^2_{0.010}$	$\chi^2_{0.005}$
1	0.00004	0.00016	0.00098	0.00393	0.0158	0.455	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	1.386	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	2.366	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	3.357	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.61	4.251	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	5.35	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.83	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.43	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	51.17	60.39	64.28	79.33	98.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	89.33	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	99.33	118.5	124.3	129.6	135.8	140.2

来源：Table Ⅳ of Ronald A. Fisher and Frank Yates, Statistical Tables for Biological, Agricultural and Medical Research(6th ed.), Longman(朗文)Group Ltd., London, 1974.(早期由 Edinburgh 的 Oliver & Boyd 有限公司出版). 由 Pearson Education Limited 授权翻印.

表 A.8 F 分布的临界值

本表给出 F 分布的临界值 f_{α,ν_1,ν_2}，表的使用说明见 17.19 节.

ν_2	α	ν_1 1	2	3	4	5	6	7	8	9
	0.100	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86
	0.050	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
1	0.025	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3
	0.010	4052.0	4999.5	5403.0	5625.0	5764.0	5859.0	5928.0	5982.0	6022.0
	0.005	16211	20000	21615	22500	23056	23437	23715	23925	24091
	0.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	0.050	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38
2	0.025	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39
	0.010	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39
	0.005	198.5	199.0	199.2	199.2	199.3	199.3	199.4	199.4	199.4
	0.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	0.050	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
3	0.025	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47
	0.010	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35
	0.005	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88
	0.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	0.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
4	0.025	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90
	0.010	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66
	0.005	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14
	0.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	0.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
5	0.025	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68
	0.010	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16
	0.005	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77
	0.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	0.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
6	0.025	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52
	0.010	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	0.005	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39
	0.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	0.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
7	0.025	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82
	0.010	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	0.005	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51
	0.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	0.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
8	0.025	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36
	0.010	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	0.005	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34
	0.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	0.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
9	0.025	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03
	0.010	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	0.005	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54

续表

ν_2	α	ν_1									
		10	12	15	20	24	30	40	60	120	∞
1	0.100	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
	0.050	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
	0.025	968.6	976.7	984.9	993.1	997.2	1001	1006	1010	1014	1018
	0.010	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
	0.005	24224	24426	24630	24836	24940	25044	25148	25253	25359	25465
2	0.100	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
	0.050	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
	0.025	39.40	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49	39.50
	0.010	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
	0.005	199.4	199.4	199.4	199.4	199.5	199.5	199.5	199.5	199.5	199.5
3	0.100	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
	0.050	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
	0.025	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95	13.90
	0.010	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
	0.005	43.69	43.39	43.08	42.78	42.62	42.47	42.31	42.15	41.99	41.83
4	0.100	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
	0.050	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
	0.025	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31	8.26
	0.010	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
	0.005	20.97	20.70	20.44	20.17	20.03	19.89	19.75	19.61	19.47	19.32
5	0.100	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
	0.050	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
	0.025	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07	6.02
	0.010	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
	0.005	13.62	13.38	13.15	12.90	12.78	12.66	12.53	12.40	12.27	12.14
6	0.100	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
	0.050	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
	0.025	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.90	4.85
	0.010	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
	0.005	10.25	10.03	9.81	9.59	9.47	9.36	9.24	9.12	9.00	8.88
7	0.100	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
	0.050	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
	0.025	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.25	4.20	4.14
	0.010	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
	0.005	8.38	8.18	7.97	7.75	7.65	7.53	7.42	7.31	7.19	7.08
8	0.100	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
	0.050	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
	0.025	4.30	4.20	4.10	4.00	3.95	3.89	3.84	3.78	3.73	3.67
	0.010	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
	0.005	7.21	7.01	6.81	6.61	6.50	6.40	6.29	6.18	6.06	5.95
9	0.100	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
	0.050	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
	0.025	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39	3.33
	0.010	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
	0.005	6.42	6.23	6.03	5.83	5.73	5.62	5.52	5.41	5.30	5.19

续表

ν_2	α	ν_1 = 1	2	3	4	5	6	7	8	9
10	0.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	0.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	0.025	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78
	0.010	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	0.005	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97
11	0.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	0.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	0.025	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59
	0.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	0.005	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54
12	0.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	0.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	0.025	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44
	0.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	0.005	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20
13	0.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	0.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	0.025	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31
	0.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	0.005	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94
14	0.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	0.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	0.025	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21
	0.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	0.005	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72
15	0.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	0.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	0.025	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12
	0.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	0.005	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54
16	0.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	0.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	0.025	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05
	0.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	0.005	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38
17	0.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	0.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	0.025	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98
	0.010	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	0.005	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25
18	0.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	0.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	0.025	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93
	0.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	0.005	10.22	7.21	6.03	5.37	4.96	4.66	4.44	4.28	4.14

续表

ν_2	α	ν_1: 10	12	15	20	24	30	40	60	120	∞
10	0.100	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
	0.050	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
	0.025	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.20	3.14	3.08
	0.010	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
	0.005	5.85	5.66	5.47	5.27	5.17	5.07	4.97	4.86	4.75	4.64
11	0.100	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
	0.050	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
	0.025	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.00	2.94	2.88
	0.010	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
	0.005	5.42	5.24	5.05	4.86	4.76	4.65	4.55	4.44	4.34	4.23
12	0.100	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
	0.050	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
	0.025	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79	2.72
	0.010	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
	0.005	5.09	4.91	4.72	4.53	4.43	4.33	4.23	4.12	4.01	3.90
13	0.100	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
	0.050	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
	0.025	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66	2.60
	0.010	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
	0.005	4.82	4.64	4.46	4.27	4.17	4.07	3.97	3.87	3.76	3.65
14	0.100	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	.83	1.80
	0.050	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
	0.025	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55	2.49
	0.010	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
	0.005	4.60	4.43	4.25	4.06	3.96	3.86	3.76	3.66	3.55	3.44
15	0.100	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	..79	1.76
	0.050	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
	0.025	3.06	2.96	2.86	2.76	2.70	2.64	2.59	2.52	2.46	2.40
	0.010	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
	0.005	4.42	4.25	4.07	3.88	3.79	3.69	3.58	3.48	3.37	3.26
16	0.100	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	.75	1.72
	0.050	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
	0.025	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38	2.32
	0.010	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
	0.005	4.27	4.10	3.92	3.73	3.64	3.54	3.44	3.33	3.22	3.11
17	0.100	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	.72	1.69
	0.050	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
	0.025	2.92	2.82	2.72	2.62	2.56	2.50	2.44	2.38	2.32	2.25
	0.010	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
	0.005	4.14	3.97	3.79	3.61	3.51	3.41	3.31	3.21	3.10	2.98
18	0.100	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	..69	1.66
	0.050	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	.97	1.92
	0.025	2.87	2.77	2.67	2.56	2.50	2.44	2.38	2.32	2.26	2.19
	0.010	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
	0.005	4.03	3.86	3.68	3.50	3.40	3.30	3.20	3.10	2.99	2.87

续表

ν_2	α	ν_1 = 1	2	3	4	5	6	7	8	9
19	0.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.96
	0.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	0.025	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88
	0.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	0.005	10.07	7.09	5.92	5.27	4.85	4.56	4.34	4.18	4.04
20	0.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	0.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	0.025	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84
	0.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	0.005	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96
21	0.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
	0.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	0.025	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80
	0.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	0.005	9.83	6.89	5.73	5.09	4.68	4.39	4.18	4.01	3.88
22	0.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	0.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	0.025	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76
	0.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	0.005	9.73	6.81	5.65	5.02	4.61	4.32	4.11	3.94	3.81
23	0.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	0.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	0.025	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73
	0.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	0.005	9.63	6.73	5.58	4.95	4.54	4.26	4.05	3.88	3.75
24	0.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	0.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	0.025	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70
	0.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	0.005	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69
25	0.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	0.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	0.025	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68
	0.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	0.005	9.48	6.60	5.46	4.84	4.43	4.15	3.94	3.78	3.64
26	0.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	0.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	0.025	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65
	0.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	0.005	9.41	6.54	5.41	4.79	4.38	4.10	3.89	3.73	3.60
27	0.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	0.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	0.025	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63
	0.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	0.005	9.34	6.49	5.36	4.74	4.34	4.06	3.85	3.69	3.56

续表

ν_2	α	ν_1 10	12	15	20	24	30	40	60	120	∞
19	0.100	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
	0.050	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
	0.025	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.20	2.13
	0.010	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
	0.005	3.93	3.76	3.59	3.40	3.31	3.21	3.11	3.00	2.89	2.78
20	0.100	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
	0.050	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
	0.025	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16	2.09
	0.010	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
	0.005	3.85	3.68	3.50	3.32	3.22	3.12	3.02	2.92	2.81	2.69
21	0.100	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
	0.050	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
	0.025	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11	2.04
	0.010	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
	0.005	3.77	3.60	3.43	3.24	3.15	3.05	2.95	2.84	2.73	2.61
22	0.100	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
	0.050	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
	0.025	2.70	2.60	2.50	2.39	2.33	2.27	2.21	2.14	2.08	2.00
	0.010	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
	0.005	3.70	3.54	3.36	3.18	3.08	2.98	2.88	2.77	2.66	2.55
23	0.100	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
	0.050	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
	0.025	2.67	2.57	2.47	2.36	2.30	2.24	2.18	2.11	2.04	1.97
	0.010	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
	0.005	3.64	3.47	3.30	3.12	3.02	2.92	2.82	2.71	2.60	2.48
24	0.100	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
	0.050	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
	0.025	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01	1.94
	0.010	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
	0.005	3.59	3.42	3.25	3.06	2.97	2.87	2.77	2.66	2.55	2.43
25	0.100	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
	0.050	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
	0.025	2.61	2.51	2.41	2.30	2.24	2.18	2.12	2.05	1.98	1.91
	0.010	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
	0.005	3.54	3.37	3.20	3.01	2.92	2.82	2.72	2.61	2.50	2.38
26	0.100	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
	0.050	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
	0.025	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95	1.88
	0.010	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
	0.005	3.49	3.33	3.15	2.97	2.87	2.77	2.67	2.56	2.45	2.33
27	0.100	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
	0.050	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
	0.025	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.00	1.93	1.85
	0.010	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
	0.005	3.45	3.28	3.11	2.93	2.83	2.73	2.63	2.52	2.41	2.29

续表

ν_2	α	ν_1 1	2	3	4	5	6	7	8	9
28	0.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	0.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	0.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	0.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	0.005	9.28	6.44	5.32	4.70	4.30	4.02	3.81	3.65	3.52
29	0.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	0.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	0.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	0.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	0.005	9.23	6.40	5.28	4.66	4.26	3.98	3.77	3.61	3.48
30	0.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	0.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	0.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	0.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	0.005	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45
40	0.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	0.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	0.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	0.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	0.005	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22
60	0.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	0.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	0.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	0.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	0.005	8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01
120	0.100	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68
	0.050	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
	0.025	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22
	0.010	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56
	0.005	8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81
∞	0.100	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	0.050	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
	0.025	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11
	0.010	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	0.005	7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62

续表

ν_2	α	ν_1 10	12	15	20	24	30	40	60	120	∞
28	0.100	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
	0.050	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
	0.025	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91	1.83
	0.010	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
	0.005	3.41	3.25	3.07	2.89	2.79	2.69	2.59	2.48	2.37	2.25
29	0.100	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
	0.050	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
	0.025	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89	1.81
	0.010	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
	0.005	3.38	3.21	3.04	2.86	2.76	2.66	2.56	2.45	2.33	2.21
30	0.100	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
	0.050	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
	0.025	2.51	2.41	2.31	2.20	2.14	2.07	2.01	1.94	1.87	1.79
	0.010	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
	0.005	3.34	3.18	3.01	2.82	2.73	2.63	2.52	2.42	2.30	2.18
40	0.100	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
	0.050	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
	0.025	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.80	1.72	1.64
	0.010	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
	0.005	3.12	2.95	2.78	2.60	2.50	2.40	2.30	2.18	2.06	1.93
60	0.100	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
	0.050	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	.47	1.39
	0.025	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	.58	1.48
	0.010	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	.73	1.60
	0.005	2.90	2.74	2.57	2.39	2.29	2.19	2.08	1.96	1.83	1.69
120	0.100	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
	0.050	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
	0.025	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43	1.31
	0.010	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
	0.005	2.71	2.54	2.37	2.19	2.09	1.98	1.87	1.75	1.61	1.43
∞	0.100	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00
	0.050	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00
	0.025	2.05	1.94	1.83	1.71	1.64	1.57	1.48	1.39	1.27	1.00
	0.010	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00
	0.005	2.52	2.36	2.19	2.00	1.90	1.79	1.67	1.53	1.36	1.00

来源:M Merrington and C. M. Thompson, "Tables of percentage points of the inverted beta(F)distribution", Biometrika, vol. 33(1943), 以及 Table 18 of E. S. Pearson and H. O. Hartley, Biometrika Tables for Statisticians, vol. 1, Cambridge University Press, Cambridge, 1954. 由 Biometrika 信托代表授权翻印.

表 A.9　最小显著的学生化极差 r_p

	$\alpha=0.05$						$\alpha=0.01$				
	p						p				
ν	2	3	4	5	6	ν	2	3	4	5	6
1	17.97	17.97	17.97	17.97	17.97	1	90.03	90.03	90.03	90.03	90.03
2	6.085	6.085	6.085	6.085	6.085	2	14.04	14.04	14.04	14.04	14.04
3	4.501	4.516	4.516	4.516	4.516	3	8.261	8.321	8.321	8.321	8.321
4	3.927	4.013	4.033	4.033	4.033	4	6.512	6.677	6.740	6.756	6.756
5	3.635	3.749	3.797	3.814	3.814	5	5.702	5.893	5.898	6.040	6.065
6	3.461	3.587	3.649	3.680	3.694	6	5.243	5.439	5.549	5.614	5.655
7	3.344	3.477	3.548	3.588	3.611	7	4.949	5.145	5.260	5.334	5.383
8	3.261	3.399	3.475	3.521	3.549	8	4.746	4.939	5.057	5.135	5.189
9	3.199	3.339	3.420	3.470	3.502	9	4.596	4.787	4.906	4.986	5.043
10	3.151	3.293	3.376	3.430	3.465	10	4.482	4.671	4.790	4.871	4.931
11	3.113	3.256	3.342	3.397	3.435	11	4.392	4.579	4.697	4.780	4.841
12	3.082	3.225	3.313	3.370	3.410	12	4.320	4.504	4.622	4.706	4.767
13	3.055	3.200	3.289	3.348	3.389	13	4.260	4.442	4.560	4.644	4.706
14	3.033	3.178	3.268	3.329	3.372	14	4.210	4.391	4.508	4.591	4.654
15	3.014	3.160	3.250	3.312	3.356	15	4.168	4.347	4.463	4.547	4.610
16	2.998	3.144	3.235	3.298	3.343	16	4.131	4.309	4.425	4.509	4.572
17	2.984	3.130	3.222	3.285	3.331	17	4.099	4.275	4.391	4.475	4.539
18	2.971	3.118	3.210	3.274	3.321	18	4.071	4.246	4.362	4.445	4.509
19	2.960	3.107	3.199	3.264	3.311	19	4.046	4.220	4.335	4.419	4.483
20	2.950	3.097	3.190	3.255	3.303	20	4.024	4.197	4.312	4.395	4.459
24	2.919	3.066	3.160	3.226	3.276	24	3.956	4.126	4.239	4.322	4.386
30	2.888	3.035	3.131	3.199	3.250	30	3.889	4.056	4.168	4.250	4.314
40	2.858	3.006	3.102	3.171	3.224	40	3.825	3.988	4.098	4.180	4.244
60	2.829	2.976	3.073	3.143	3.198	60	3.762	3.922	4.031	4.111	4.174
120	2.800	2.947	3.045	3.116	3.172	120	3.702	3.858	3.965	4.044	4.107
∞	2.772	2.918	3.017	3.089	3.146	∞	3.643	3.796	3.900	3.978	4.040

来源：H. L. Harter, "Critical values for Duncan's new multiple range test," Biometrics, vol. 16(1960). 由国际生物统计学会(IBS)授权翻印.

表 A.10　r 到 z_r 的变换

本表给出 Pearson 乘积矩相关系数 r 的 z_r 值,表的使用说明见 19.14 节.

r	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.000	0.0000	0.0010	0.0020	0.0030	0.0040	0.0050	0.0060	0.0070	0.0080	0.0090
0.010	0.0100	0.0110	0.0120	0.0130	0.0140	0.0150	0.0160	0.0170	0.0180	0.0190
0.020	0.0200	0.0210	0.0220	0.0230	0.0240	0.0250	0.0260	0.0270	0.0280	0.0290
0.030	0.0300	0.0310	0.0320	0.0330	0.0340	0.0350	0.0360	0.0370	0.0380	0.0390
0.040	0.0400	0.0410	0.0420	0.0430	0.0440	0.0450	0.0460	0.0470	0.0480	0.0490
0.050	0.0501	0.0511	0.0521	0.0531	0.0541	0.0551	0.0561	0.0571	0.0581	0.0591
0.060	0.0601	0.0611	0.0621	0.0631	0.0641	0.0651	0.0661	0.0671	0.0681	0.0691
0.070	0.0701	0.0711	0.0721	0.0731	0.0741	0.0751	0.0761	0.0771	0.0782	0.0792
0.080	0.0802	0.0812	0.0822	0.0832	0.0842	0.0852	0.0862	0.0872	0.0882	0.0892
0.090	0.0902	0.0912	0.0922	0.0933	0.0943	0.0953	0.0963	0.0973	0.0983	0.0993
0.100	0.1003	0.1013	0.1024	0.1034	0.1044	0.1054	0.1064	0.1074	0.1084	0.1094
0.110	0.1105	0.1115	0.1125	0.1135	0.1145	0.1155	0.1165	0.1175	0.1185	0.1195
0.120	0.1206	0.1216	0.1226	0.1236	0.1246	0.1257	0.1267	0.1277	0.1287	0.1297
0.130	0.1308	0.1318	0.1328	0.1338	0.1348	0.1358	0.1368	0.1379	0.1389	0.1399
0.140	0.1409	0.1419	0.1430	0.1440	0.1450	0.1460	0.1470	0.1481	0.1491	0.1501
0.150	0.1511	0.1522	0.1532	0.1542	0.1552	0.1563	0.1573	0.1583	0.1593	0.1604
0.160	0.1614	0.1624	0.1634	0.1644	0.1655	0.1665	0.1676	0.1686	0.1696	0.1706
0.170	0.1717	0.1727	0.1737	0.1748	0.1758	0.1768	0.1779	0.1789	0.1799	0.1810
0.180	0.1820	0.1830	0.1841	0.1851	0.1861	0.1872	0.1882	0.1892	0.1903	0.1913
0.190	0.1923	0.1934	0.1944	0.1954	0.1965	0.1975	0.1986	0.1996	0.2007	0.2017
0.200	0.2027	0.2038	0.2048	0.2059	0.2069	0.2079	0.2090	0.2100	0.2 11	0.2121
0.210	0.2132	0.2142	0.2153	0.2163	0.2174	0.2184	0.2194	0.2205	0.2215	0.2226
0.220	0.2237	0.2247	0.2258	0.2268	0.2279	0.2289	0.2300	0.2310	0.2321	0.2331
0.230	0.2342	0.2353	0.2363	0.2374	0.2384	0.2395	0.2405	0.2416	0.2427	0.2437
0.240	0.2448	0.2458	0.2469	0.2480	0.2490	0.2501	0.2511	0.2522	0.2533	0.2543
0.250	0.2554	0.2565	0.2575	0.2586	0.2597	0.2608	0.2618	0.2629	0.2640	0.2650
0.260	0.2661	0.2672	0.2682	0.2693	0.2704	0.2715	0.2726	0.2736	0.2747	0.2758
0.270	0.2769	0.2779	0.2790	0.2801	0.2812	0.2823	0.2833	0.2844	0.2855	0.2866
0.280	0.2877	0.2888	0.2898	0.2909	0.2920	0.2931	0.2942	0.2953	0.2964	0.2975
0.290	0.2986	0.2997	0.3008	0.3019	0.3029	0.3040	0.3051	0.3062	0.3073	0.3084
0.300	0.3095	0.3106	0.3117	0.3128	0.3139	0.3150	0.3161	0.3172	0.3183	0.3195
0.310	0.3206	0.3217	0.3228	0.3239	0.3250	0.3261	0.3272	0.3283	0.3294	0.3305
0.320	0.3317	0.3328	0.3339	0.3350	0.3361	0.3372	0.3384	0.3395	0.3406	0.3417
0.330	0.3428	0.3439	0.3451	0.3462	0.3473	0.3484	0.3496	0.3507	0.3518	0.3530
0.340	0.3541	0.3552	0.3564	0.3575	0.3586	0.3597	0.3609	0.3620	0.3632	0.3643
0.350	0.3654	0.3666	0.3677	0.3689	0.3700	0.3712	0.3723	0.3734	0.3746	0.3757
0.360	0.3769	0.3780	0.3792	0.3803	0.3815	0.3826	0.3838	0.3850	0.3861	0.3873
0.370	0.3884	0.3896	0.3907	0.3919	0.3931	0.3942	0.3954	0.3966	0.3977	0.3989
0.380	0.4001	0.4012	0.4024	0.4036	0.4047	0.4059	0.4071	0.4083	0.4094	0.4106
0.390	0.4118	0.4130	0.4142	0.4153	0.4165	0.4177	0.4189	0.4201	0.4213	0.4225
0.400	0.4236	0.4248	0.4260	0.4272	0.4284	0.4296	0.4308	0.4320	0.4332	0.4344
0.410	0.4356	0.4368	0.4380	0.4392	0.4404	0.4416	0.4429	0.4441	0.4453	0.4465
0.420	0.4477	0.4489	0.4501	0.4513	0.4526	0.4538	0.4550	0.4562	0.4574	0.4587
0.430	0.4599	0.4611	0.4623	0.4636	0.4648	0.4660	0.4673	0.4685	0.4697	0.4710
0.440	0.4722	0.4735	0.4747	0.4760	0.4772	0.4784	0.4797	0.4809	0.4822	0.4835
0.450	0.4847	0.4860	0.4872	0.4885	0.4897	0.4910	0.4923	0.4935	0.4948	0.4961
0.460	0.4973	0.4986	0.4999	0.5011	0.5024	0.5037	0.5049	0.5062	0.5075	0.5088
0.470	0.5101	0.5114	0.5126	0.5139	0.5152	0.5165	0.5178	0.5191	0.5204	0.5217
0.480	0.5230	0.5243	0.5256	0.5279	0.5282	0.5295	0.5308	0.5321	0.5334	0.5347
0.490	0.5361	0.5374	0.5387	0.5400	0.5413	0.5427	0.5440	0.5453	0.5466	0.5480

续表

r	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.500	0.5493	0.5506	0.5520	0.5533	0.5547	0.5560	0.5573	0.5587	0.5600	0.5614
0.510	0.5627	0.5641	0.5654	0.5668	0.5681	0.5695	0 5709	0.5722	0.5736	0.5750
0.520	0.5763	0.5777	0.5791	0.5805	0.5818	0.5832	0.5846	0.5860	0.5847	0.5888
0.530	0.5901	0.5915	0.5929	0.5943	0.5957	0.5971	0.5985	0.5999	0.6013	0.6027
0.540	0.6042	0.6056	0.6070	0.6084	0.6098	0.6112	0.6127	0.6141	0.6155	0.6170
0.550	0.6184	0.6198	0.6213	0.6227	0.6241	0.6256	0.6270	0.6285	0.6299	0.6314
0.560	0.6328	0.6343	0.6358	0.6372	0.6387	0.6401	0.6416	0.6431	0.6446	0.6460
0.570	0.6475	0.6490	0.6505	0.6520	0.6535	0.6550	0.6565	0.6579	0.6594	0.6610
0.580	0.6625	0.6640	0.6655	0.6670	0.6685	0.6700	0.6715	0.6731	0.6746	0.6761
0.590	0.6777	0.6792	0.6807	0.6823	0.6838	0.6854	0.6869	0.6885	0.6900	0.6916
0.600	0.6931	0.6947	0.6963	0.6978	0.6994	0.7010	0.7026	0.7042	0.7057	0.7073
0.610	0.7089	0.7105	0.7121	0.7137	0.7153	0.7169	0.7185	0.7201	0.7218	0.7234
0.620	0.7250	0.7266	0.7283	0.7299	0.7315	0.7332	0.7348	0.7364	0.7381	0.7398
0.630	0.7414	0.7431	0.7447	0.7464	0.7481	0.7497	0.7514	0.7531	0.7548	0.7565
0.640	0.7582	0.7599	0.7616	0.7633	0.7650	0.7667	0.7684	0.7701	0.7718	0.7736
0.650	0.7753	0.7770	0.7788	0.7805	0.7823	0.7840	0.7858	0.7875	0.7893	0.7910
0.660	0.7928	0.7946	0.7964	0.7981	0.7999	0.8017	0.8035	0.8053	0.8071	0.8089
0.670	0.8107	0.8126	0.8144	0.8162	0.8180	0.8199	0.8217	0.8236	0.8254	0.8273
0.680	0.8291	0.8310	0.8328	0.8347	0.8366	0.8385	0.8404	0.8423	0.8442	0.8461
0.690	0.8480	0.8499	0.8518	0.8537	0.8556	0.8576	0.8595	0.8614	0.8634	0.8653
0.700	0.8673	0.8693	0.8712	0.8732	0.8752	0.8772	0.8792	0.8812	0.8832	0.8852
0.710	0.8872	0.8892	0.8912	0.8933	0.8953	0.8973	0.8994	0.9014	0.9035	0.9056
0.720	0.9076	0.9097	0.9118	0.9139	0.9160	0.9181	0.9202	0.9223	0.9245	0.9266
0.730	0.9287	0.9309	0.9330	0.9352	0.9373	0.9395	0.9417	0.9439	0.9461	0.9483
0.740	0.9505	0.9527	0.9549	0.9571	0.9594	0.9616	0.9639	0.9661	0.9684	0.9707
0.750	0.9730	0.9752	0.9775	0.9799	0.9822	0.9845	0.9868	0.9892	0.9915	0.9939
0.760	0.9962	0.9986	1.0010	1.0034	1.0058	1.0082	1.0106	1.0130	1.0154	1.0179
0.770	1.0203	1.0228	1.0253	1.0277	1.0302	1.0327	1.0352	1.0378	1.0403	1.0428
0.780	1.0454	1.0479	1.0505	1.0531	1.0557	1.0583	1.0609	1.0635	1.0661	1.0688
0.790	1.0714	1.0741	1.0768	1.0795	1.0822	1.0849	1.0876	1.0903	1.0931	1.0958
0.800	1.0986	1.1014	1.1041	1.1070	1.1098	1.1127	1.1155	1.1184	1.1212	1.1241
0.810	1.1270	1.1299	1.1329	1.1358	1.1388	1.1417	1.1447	1.1477	1.1507	1.1538
0.820	1.1568	1.1599	1.1630	1.1660	1.1692	1.1723	1.1754	1.1786	1.1817	1.1849
0.830	1.1870	1.1913	1.1946	1.1979	1.2011	1.2044	1.2077	1.2111	1.2144	1.2178
0.840	1.2212	1.2246	1.2280	1.2315	1.2349	1 2384	1.2419	1.2454	1.2490	1.2526
0.850	1.2561	1.2598	1.2634	1.2670	1.2708	1.2744	1.2782	1.2819	1.2857	1.2895
0.860	1.2934	1.2972	1.3011	1.3050	1.3089	1.3129	1.3168	1.3209	1.3249	1.3290
0.870	1.3331	1.3372	1.3414	1.3456	1.3498	1.3540	1.3583	1.3626	1.3670	1.3714
0.880	1.3758	1.3802	1.3847	1.3892	1.3938	1.3984	1.4030	1.4077	1.4124	1.4171
0 890	1.4219	1.4268	1.4316	1.4366	1.4415	1.4465	1.4516	1.4566	1.4618	1.4670
0.900	1.4722	1.4775	1.4828	1.4883	1.4937	1.4992	1.5047	1.5103	1.5160	1.5217
0.910	1.5275	1.5334	1.5393	1.5453	1.5513	1.5574	1.5636	1.5698	1.5762	1.5825
0.920	1.5890	1.5956	1.6022	1.6089	1.6157	1.6226	1.6296	1.6366	1.6438	1.6510
0.930	1.6584	1.6659	1.6734	1.6811	1.6888	1.6967	1.7047	1.7129	1.7211	1.7295
0.940	1.7380	1.7467	1.7555	1.7646	1.7736	1.7828	1.7923	1.8019	1.8117	1.8216
0.950	1.8318	1.8421	1.8527	1.8635	1.8745	1.8857	1.8972	1.9090	1.9210	1.9333
0.960	1.9459	1.9588	1.9721	1.9857	1.9996	2.0140	2.0287	2.0439	2.0595	2.0756
0 970	2.0923	2.1095	2.1273	2.1457	2.1649	2.1847	2.2054	2.2269	2.2494	2.2729
0.980	2.2976	2.3223	2.3507	2.3796	2.4101	2.4426	2.4774	2.5147	2.5550	2.5988
0.990	2.6467	2.6996	2.7587	2.8257	2.9031	2.9945	3.1063	3.2504	3.4534	3.8002

r	z_r
0.9999	4.95172
0.99999	6.10303

来源：Albert E. Waugh, Statistical Tables and Problems, McGraw-Hill, New York, 1952, 经 Herbert M. Blalock Jr. 修改, Social Statistics, McGraw-Hill, New York, 1979 由 McGraw-Hill 公司授权翻印.

表 A.11　Pearson 乘积矩相关系数 r 的临界值

本表给出 r 分布显著性水平为 α 的临界值,表的使用说明见 19.15 节.

	Levels of significance for a one-tailed test					Levels of significance for a one-tailed test			
	0.05	0.025	0.01	0.005		0.05	0.025	0.11	0.05
	Levels of significance for a two-tailed test					Levels of significance for a two-tailed test			
ν	0.10	0.05	0.02	0.01	ν	0.10	0.05	0.12	0.01
1	0.988	0.997	0.9995	0.9999	24	0.330	0.388	0.453	0.496
2	0.900	0.950	0.980	0.990	26	0.317	0.374	0.437	0.479
3	0.805	0.878	0.934	0.959	28	0.306	0.361	0.423	0.463
4	0.729	0.811	0.882	0.917	30	0.296	0.349	0.409	0.449
5	0.669	0.755	0.833	0.875	35	0.275	0.325	0.381	0.418
6	0.622	0.707	0.789	0.834	40	0.257	0.304	0.358	0.393
7	0.582	0.666	0.750	0.798	45	0.243	0.288	0.338	0.372
8	0.549	0.632	0.716	0.765	50	0.231	0.273	0.322	0.354
9	0.521	0.602	0.685	0.735	55	0.220	0.261	0.307	0.339
10	0.497	0.576	0.658	0.708	60	0.211	0.250	0.295	0.325
11	0.476	0.553	0.634	0.684	70	0.195	0.232	0.274	0.302
12	0.458	0.532	0.612	0.661	80	0.183	0.217	0.256	0.283
13	0.441	0.514	0.592	0.641	90	0.173	0.205	0.242	0.267
14	0.426	0.497	0.574	0.623	100	0.164	0.195	0.230	0.254
15	0.412	0.482	0.558	0.606	120	0.150	0.178	0.2[illegible]0	0.232
16	0.400	0.468	0.542	0.590	150	0.134	0.159	0.189	0.208
17	0.389	0.456	0.529	0.575	200	0.116	0.138	0.164	0.181
18	0.378	0.444	0.516	0.561	300	0.095	0.113	0.134	0.148
19	0.369	0.433	0.503	0.549	400	0.082	0.098	0.1[illegible]6	0.128
20	0.360	0.423	0.492	0.537	500	0.073	0.088	0.104	0.115
22	0.344	0.404	0.472	0.515	1000	0.052	0.062	0.073	0.081

来源:Table VI of Ronald A. Fisher and Frank Yates, Statistical Tables for Biological, Agricultural and Medical Research(6th ed.), Longman(朗文)Group Ltd., London, 1974.(早期由 Edinburgh 的 Oliver & Boyd 有限公司出版.)由 Pearson Education Limited 授权翻印. 经 E. W. Minium and R. B. Clarke 修改, Elements of Statistical Reasoning, John Wiley & Sons, Inc., 1982 由 John Wiley & Sons, Inc. 授权翻印.

表 A.12　Wilcoxon W 的临界值

本表给出 $n\leqslant 50$ 时显著性水平为 α 的检验统计量 W 的临界值，表的使用说明见 20.7 节.

单尾	双尾	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
$\alpha=0.05$	$\alpha=0.10$	1	2	4	6	8	11
$\alpha=0.025$	$\alpha=0.05$		1	2	4	6	8
$\alpha=0.01$	$\alpha=0.02$			0	2	3	5
$\alpha=0.005$	$\alpha=0.01$				0	2	3
单尾	双尾	$n=11$	$n=12$	$n=13$	$n=14$	$n=15$	$n=16$
$\alpha=0.05$	$\alpha=0.10$	14	17	21	26	30	36
$\alpha=0.025$	$\alpha=0.05$	11	14	17	21	25	30
$\alpha=0.01$	$\alpha=0.02$	7	10	13	16	20	24
$\alpha=0.005$	$\alpha=0.01$	5	7	10	13	16	19
单尾	双尾	$n=17$	$n=18$	$n=19$	$n=20$	$n=21$	$n=22$
$\alpha=0.05$	$\alpha=0.10$	41	47	54	60	68	75
$\alpha=0.025$	$\alpha=0.05$	35	40	46	52	59	66
$\alpha=0.01$	$\alpha=0.02$	28	33	38	43	49	56
$\alpha=0.005$	$\alpha=0.01$	23	28	32	37	43	49
单尾	双尾	$n=23$	$n=24$	$n=25$	$n=26$	$n=27$	$n=28$
$\alpha=0.05$	$\alpha=0.10$	83	92	101	110	120	130
$\alpha=0.025$	$\alpha=0.05$	73	81	90	98	107	117
$\alpha=0.01$	$\alpha=0.02$	62	69	77	85	93	102
$\alpha=0.005$	$\alpha=0.01$	55	61	68	76	84	92
单尾	双尾	$n=29$	$n=30$	$n=31$	$n=32$	$n=33$	$n=34$
$\alpha=0.05$	$\alpha=0.10$	141	152	163	175	188	201
$\alpha=0.025$	$\alpha=0.05$	127	137	148	159	171	183
$\alpha=0.01$	$\alpha=0.02$	111	120	130	141	151	162
$\alpha=0.005$	$\alpha=0.01$	100	109	118	128	138	149
单尾	双尾	$n=35$	$n=36$	$n=37$	$n=38$	$n=39$	$n=40$
$\alpha=0.05$	$\alpha=0.10$	214	228	242	256	271	287
$\alpha=0.025$	$\alpha=0.05$	195	208	222	235	250	264
$\alpha=0.01$	$\alpha=0.02$	174	186	198	211	224	238
$\alpha=0.005$	$\alpha=0.01$	160	171	183	195	208	221
单尾	双尾	$n=41$	$n=42$	$n=43$	$n=44$	$n=45$	$n=46$
$\alpha=0.05$	$\alpha=0.10$	303	319	336	353	371	389
$\alpha=0.025$	$\alpha=0.05$	279	295	311	327	344	361
$\alpha=0.01$	$\alpha=0.02$	252	267	281	297	313	329
$\alpha=0.005$	$\alpha=0.01$	234	248	262	277	292	307
单尾	双尾	$n=47$	$n=48$	$n=49$	$n=50$		
$\alpha=0.05$	$\alpha=0.10$	408	427	446	466		
$\alpha=0.025$	$\alpha=0.05$	379	397	415	434		
$\alpha=0.01$	$\alpha=0.02$	345	362	380	398		
$\alpha=0.005$	$\alpha=0.01$	323	339	356	373		

来源：W. H. Beyer(ed.), CRC Handbook of Tables for Probability and Statistics(2nd ed.), CRC Press, Inc., Boca Raton, Florida, 1968. 由 CRC Press, Inc. 授权翻印.

表 A.13　Mann-Whitney U 的临界值

本表给出小样本($n_1 \leqslant 20$, $n_2 \leqslant 20$)时显著性水平为 $\alpha = 0.05$(罗马体)与 $\alpha = 0.025$(黑体)的单尾检验及 $\alpha = 0.10$(罗马体)与 $\alpha = 0.05$(黑体)时的双尾检验的临界值. 表中的‘－’线表示对给定的 ι,不能作出决策. 表的使用说明见 20.9 节.

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–		–	–	–	–	–	–	–	–	–	–	–	–	–	–		0	0
																			–	
2	–	–	–	–	0	0	0	1	1	1	1	2	2	2	3	3	3	4	4	4
					–	–	–	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**1**	**1**	**2**	**2**	**2**	**2**
3	–	–	0	0	1	2	2	3	3	4	5	5	6	7	7	8	9	9	10	11
	–	–	–	–	**0**	**1**	**1**	**2**	**2**	**3**	**3**	**4**	**4**	**5**	**5**	**6**	**6**	**7**	**7**	**8**
4	–	–	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
	–	–	–	**0**	**1**	**2**	**3**	**4**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**11**	**12**	**13**	**13**
5	–	0	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
	–	–	**0**	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**	**17**	**18**	**19**	**20**
6	–	0	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
	–	–	**1**	**2**	**3**	**5**	**6**	**8**	**10**	**11**	**13**	**14**	**16**	**17**	**19**	**21**	**22**	**24**	**25**	**27**
7	–	0	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
	–	–	**1**	**3**	**5**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**	**28**	**30**	**32**	**34**
8	–	1	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
	–	**0**	**2**	**4**	**6**	**8**	**10**	**13**	**15**	**17**	**19**	**22**	**24**	**26**	**29**	**31**	**34**	**36**	**38**	**41**
9	–	1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
	–	**0**	**2**	**4**	**7**	**10**	**12**	**15**	**17**	**20**	**23**	**26**	**28**	**31**	**34**	**37**	**39**	**42**	**45**	**48**
10	–	1	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
	–	**0**	**3**	**5**	**8**	**11**	**14**	**17**	**20**	**23**	**26**	**29**	**33**	**36**	**39**	**42**	**45**	**48**	**52**	**55**
11	–	1	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
	–	**0**	**3**	**6**	**9**	**13**	**16**	**19**	**23**	**26**	**30**	**33**	**37**	**40**	**44**	**47**	**51**	**55**	**58**	**62**
12	–	2	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
	–	**1**	**4**	**7**	**11**	**14**	**18**	**22**	**26**	**29**	**33**	**37**	**41**	**45**	**49**	**53**	**57**	**61**	**65**	**69**
13	–	2	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
	–	**1**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**33**	**37**	**41**	**45**	**50**	**54**	**59**	**63**	**67**	**72**	**76**
14	–	2	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
	–	**1**	**5**	**9**	**13**	**17**	**22**	**26**	**31**	**36**	**40**	**45**	**50**	**55**	**59**	**64**	**67**	**74**	**78**	**83**
15	–	3	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
	–	**1**	**5**	**10**	**14**	**19**	**24**	**29**	**34**	**39**	**44**	**49**	**54**	**59**	**64**	**70**	**75**	**80**	**85**	**90**
16	–	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
	–	**1**	**6**	**11**	**15**	**21**	**26**	**31**	**37**	**42**	**47**	**53**	**59**	**64**	**70**	**75**	**81**	**86**	**92**	**98**
17	–	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
	–	**2**	**6**	**11**	**17**	**22**	**28**	**34**	**39**	**45**	**51**	**57**	**63**	**67**	**75**	**81**	**87**	**93**	**99**	**105**
18	–	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
	–	**2**	**7**	**12**	**18**	**24**	**30**	**36**	**42**	**48**	**55**	**61**	**67**	**74**	**80**	**86**	**93**	**99**	**106**	**112**
19	0	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
	–	**2**	**7**	**13**	**19**	**25**	**32**	**38**	**45**	**52**	**58**	**65**	**72**	**78**	**85**	**92**	**99**	**106**	**113**	**119**
20	0	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	–	**2**	**8**	**13**	**20**	**27**	**34**	**41**	**48**	**55**	**62**	**69**	**76**	**83**	**90**	**98**	**105**	**112**	**119**	**127**

表 A.14　Kruskal-Wallis H 的临界值

本表给出小样本情况($n\leqslant 5$)三种处理时显著性水平为 α 的临界值,表的使用说明见 20.10 节.

n_1	n_2	n_3	$\alpha=0.10$	0.05	0.01	n_1	n_2	n_3	$\alpha=0.10$	0.05	0.01
3	2	1	4.286			5	2	1	4.200	5.000	
3	2	2	4.500	4.714		5	2	2	4.373	5.160	6.533
3	3	1	4.571	5.143		5	3	1	4.018	4.960	
3	3	2	4.556	5.361		5	3	2	4.651	5.251	6.909
3	3	3	4.622	5.600	7.200	5	3	3	4.533	5.648	7.079
4	2	1	4.500			5	4	1	3.987	4.986	6.954
4	2	2	4.458	5.333		5	4	2	4.541	5.273	7.204
4	3	1	4.056	5.208		5	4	3	4.549	5.656	7.445
4	3	2	4.511	5.444	6.444	5	4	4	4.619	5.657	7.760
4	3	3	4.709	5.727	6.746	5	5	1	4.109	5.127	7.309
4	4	1	4.167	4.967	6.667	5	5	2	4.623	5.338	7.338
4	4	2	4.554	5.455	7.036	5	5	3	5.545	5.705	7.578
4	4	3	4.546	5.598	7.144	5	5	4	4.523	5.666	7.823
4	4	4	4.654	5.692	7.654	5	5	5	4.560	5.780	8.000

来源:W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," The Journal of American Statistical Association, vol. 47(1952). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n \leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.662	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.557	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n \leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.557	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n \leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n \leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.557	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n\leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15 Spearman r_s 的临界值

本表给出小样本($n \leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15 Spearman r_s 的临界值

本表给出小样本($n \leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n \leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.557	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n\leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.557	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H　Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n \leqslant 30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972). 由 The Journal of American Statistical Association 授权翻印.

表 A.15　Spearman r_s 的临界值

本表给出小样本($n\leqslant30$)时 r_s 的临界值,表的使用说明见 20.11 节.

	Levels of significance for a one-tailed test								
	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
	Levels of significance for a two-tailed test								
n	0.50	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
5	0.500	0.800	0.900	1.000	1.000				
6	0.371	0.657	0.829	0.886	0.943	1.000	1.000		
7	0.321	0.571	0.714	0.786	0.893	0.929	0.964	1.000	1.000
8	0.310	0.524	0.643	0.738	0.833	0.881	0.905	0.952	0.976
9	0.267	0.483	0.600	0.700	0.783	0.833	0.867	0.917	0.933
10	0.248	0.455	0.564	0.6648	0.745	0.794	0.830	0.879	0.903
11	0.236	0.427	0.536	0.618	0.709	0.755	0.800	0.845	0.873
12	0.224	0.406	0.503	0.587	0.671	0.727	0.776	0.825	0.860
13	0.209	0.385	0.484	0.560	0.648	0.703	0.747	0.802	0.835
14	0.200	0.367	0.464	0.538	0.622	0.675	0.723	0.776	0.811
15	0.189	0.354	0.443	0.521	0.604	0.654	0.700	0.754	0.786
16	0.182	0.341	0.429	0.503	0.582	0.635	0.679	0.732	0.765
17	0.176	0.328	0.414	0.485	0.566	0.615	0.662	0.713	0.748
18	0.170	0.317	0.401	0.472	0.550	0.600	0.643	0.695	0.728
19	0.165	0.309	0.391	0.460	0.535	0.584	0.628	0.677	0.712
20	0.161	0.299	0.380	0.447	0.520	0.570	0.612	0.652	0.696
21	0.156	0.292	0.370	0.435	0.508	0.556	0.599	0.648	0.681
22	0.152	0.284	0.361	0.425	0.496	0.544	0.586	0.634	0.667
23	0.148	0.278	0.353	0.415	0.486	0.532	0.573	0.622	0.654
24	0.144	0.271	0.344	0.406	0.476	0.521	0.562	0.610	0.642
25	0.142	0.265	0.337	0.398	0.466	0.511	0.551	0.598	0.630
26	0.138	0.259	0.331	0.390	0.457	0.501	0.541	0.587	0.619
27	0.136	0.255	0.324	0.382	0.448	0.491	0.531	0.577	0.608
28	0.133	0.250	0.317	0.375	0.440	0.483	0.522	0.567	0.598
29	0.130	0.245	0.312	0.368	0.433	0.475	0.513	0.558	0.589
30	0.128	0.240	0.306	0.362	0.425	0.467	0.504	0.549	0.580

来源:J H Zar, "Significance testing of the Spearman rank correlation coefficient," The Journal of American Statistical Association, vol. 67(1972).